

Institut für Angewandte Physik/Photophysik  
Fachrichtung Physik  
Fakultät Mathematik und Naturwissenschaften  
Technische Universität Dresden

# Das unstetige GALERKIN-Verfahren in der Nanooptik

Dissertation  
zur Erlangung des akademischen Grades  
Doctor rerum naturalium  
(Dr. rer. nat.)

vorgelegt von

Andreas Hille  
geboren am 01.05.1981 in Dresden

Dresden 2012

Eingereicht am: 12.09.2012

Verteidigt am: 21.12.2012

1. Gutachter: Prof. Dr. phil. II. Lukas M. Eng

2. Gutachter: Prof. Dr. rer. nat. Kurt Busch

# Kurzdarstellung

Die Nanooptik beschäftigt sich mit der Wechselwirkung von Licht mit Materie, deren charakteristische Dimension im Nanometer Bereich liegt. Insbesondere wenn die Materie aus Metall besteht, zeigen sich interessante, wellenlängenabhängige Unterschiede in der Stärke der Wechselwirkung. Die Ursache dafür sind die kollektiven Moden der quasifreien Ladungsträger, die Plasmonen. Obgleich sich experimentelle Methoden in den letzten Jahren stetig verbessert haben, ist es nach wie vor nur mit erheblichem Aufwand möglich, sich Einblicke in die mikroskopischen Zusammenhänge zu verschaffen. Eine Ergänzung zu den Experimenten bieten theoretische Modelle. Auf Grund der sich mit der Zeit stetig verbesserten Leistung der Rechentechnik, kommen dabei zunehmend numerische Verfahren zum Einsatz. Eines dieser Verfahren ist das Unstetige GALERKIN Verfahren, welches in dieser Arbeit auf folgende Fragestellungen der plasmonischen Nanooptik angewandt wurde:

- Bei dem unstetigen GALERKIN Verfahren werden die zu simulierenden Körper üblicherweise mittels Dreiecke und Tetraeder approximiert. Da die Geometrie der metallischen Systeme einen entscheidenden Einfluss auf die Wechselwirkung hat, wurde untersucht, inwieweit sich durch Einsatz von Elementen mit gekrümmten Flächen die Genauigkeit oder die Geschwindigkeit der Simulation steigern lässt. Es konnte gezeigt werden, dass runde Elemente die Genauigkeit bei gleicher Diskretisierung um bis zu zwei Größenordnungen steigern oder die Rechenzeit bei gleicher Genauigkeit auf ein Sechstel verkürzen können.
- Bestrahlt man Metallnanopartikel mit intensiven Laserpulsen, so strahlen diese nicht nur bei der Frequenz des eingestrahlteten Lichtes, sondern auch bei der doppelten Frequenz ab. Dieses Phänomen der Frequenzverdopplung (SHG, engl.: „Second-Harmonic-Generation“) ist unter anderem von der Form der Partikel und der Wellenlänge des Pulses abhängig. Da durchstimmbare gepulste Laser sehr teuer sind, wurde untersucht, ob sich mit Hilfe der linearen Partikelspektren Vorhersagen über die Stärke der Frequenzverdopplung machen lassen. Dabei wurde festgestellt, dass die Effizienz der Frequenzverdopplung zunimmt, wenn man die linearen Resonanzen der Partikel auf die SHG- oder Anregungswellenlänge abstimmt. Schafft man es, das plasmonische System so einzustellen, dass sowohl die Anregungswellenlänge, wie auch die SHG-Wellenlänge auf einer linearen Resonanz liegen, so kann die Effizienz der SHG weiter gesteigert werden.

# Abstract

Nanooptics is a discipline dealing with the interaction of light with matter where its characteristic dimensions are defined to be in the range of nanometers. In particular, if the matter consists of metal, i.e. conductive material, interesting wavelength dependent phenomena can be observed, which scale with the strength of the interaction. These phenomena are caused by the formation of collective modes between quasi-free charge carriers resulting in so called plasmons. Although improved experimental methods have evolved over the last few years, insight into the microscopic relationship between light and matter is only achievable with high effort. Supplemental information to experimental findings can be drawn from theoretical models. Due to the constantly improving computational power, numerical methods are progressively more employed. One of these methods is the discontinuous GALERKIN method, which was applied to the following problems in plasmonic nanooptics:

- Within the discontinuous GALERKIN method the simulated objects are usually approximated by triangles or tetrahedrons. Since the geometry of conductive systems has a major impact on the interaction between light and matter, the usability of elements with curved surfaces for the discretisation of the space has been investigated with respect to accuracy and speed of the simulation. In this work, it could be shown that curved elements improve the simulations precision up to two orders of magnitude with the same amount of discretisation compared to linear elements. Related to speed, it has been found that the computational time is reduced by a factor of 6 with a comparable simulation accuracy.
- By irradiating metallic nanoparticles with high power laser pulses these particles do not only emit light of the same frequency as the incident electromagnetic wave, but also with the doubled frequency (SHG, second harmonic generation). Among other things, this phenomenon of frequency doubling mainly depends on the geometry of the particle and the wavelength of the pulse. Since tunable pulsed laser sources are very expensive, it has been theoretically investigated if the strength of the frequency doubling can be deduced from the particles linear spectra. By this, it has been discovered that the efficiency of frequency doubling can be improved by adjusting the linear resonances of the particle to the SHG or excitation wavelength. The SHG efficiency can be increased even further, if the plasmonic system is tuned to a point where both the excitation and the SHG wavelength correspond to a linear resonance of the nanoparticle.



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Grundlagen</b>	<b>3</b>
2.1. Die MAXWELL - Gleichungen . . . . .	3
2.1.1. MAXWELL-Gleichung in niedrigeren Dimensionen . . . . .	5
2.1.2. Die zeitharmonischen MAXWELL-Gleichungen . . . . .	6
2.1.3. Stetigkeit der Felder . . . . .	8
2.1.4. Lösungen der MAXWELL-Gleichungen . . . . .	9
2.2. EULER - Gleichung . . . . .	19
2.2.1. Innere Energie und Druck für nichtviskose Gase . . . . .	19
2.2.2. Innere Energie und Druck für Elektronengase . . . . .	20
2.3. Materialmodelle . . . . .	22
2.3.1. DRUDE-LORENTZ Modell . . . . .	22
2.3.2. PML - Antireflexschicht . . . . .	26
2.3.3. Hydrodynamisches Elektronengas . . . . .	29
2.4. Plasmonik . . . . .	31
2.4.1. Filme . . . . .	31
2.4.2. Partikel . . . . .	34
2.5. Frequenzverdopplung in Metallen . . . . .	38
2.6. Numerische Verfahren . . . . .	42
2.6.1. Einteilung der numerischen Verfahren . . . . .	42
2.6.2. Aufbau bei DG . . . . .	45
2.6.3. Anregungsfelder . . . . .	46
2.6.4. Diskrete FOURIER Transformation . . . . .	48
2.6.5. Dimensionslose Gleichungen . . . . .	50
<b>3. Unstetiges GALERKIN-Verfahren</b>	<b>53</b>
3.1. Allgemeine Herleitung . . . . .	53
3.1.1. Basisfunktionen . . . . .	59

3.1.2.	Lokale Operatoren . . . . .	68
3.1.3.	Zeitschrittverfahren . . . . .	71
3.2.	MAXWELL-Gleichung in DG . . . . .	74
3.2.1.	Randbedingungen . . . . .	76
3.2.2.	Implementationstests . . . . .	77
3.3.	EULER-Gleichung in DG . . . . .	86
3.3.1.	Randbedingungen . . . . .	89
3.3.2.	Konvergenztests . . . . .	90
<b>4.</b>	<b>Gekrümmte Elemente</b>	<b>93</b>
4.1.	Allgemeines Verfahren . . . . .	94
4.2.	Gekrümmte Elemente in 2D . . . . .	98
4.3.	Gekrümmte Elemente in 3D . . . . .	101
<b>5.</b>	<b>Ergebnisse</b>	<b>105</b>
5.1.	Plasmonische Nanostrukturen im 2D . . . . .	105
5.1.1.	Partikel . . . . .	105
5.1.2.	V-Graben . . . . .	115
5.2.	3D-Plasmonische Nanostrukturen . . . . .	120
5.2.1.	Partikel . . . . .	120
5.2.2.	Flache IR-Antennen . . . . .	127
5.2.3.	s-SNOM . . . . .	132
5.3.	3D - SHG - Spektren . . . . .	140
5.3.1.	Einzelnes Partikel . . . . .	141
5.3.2.	Dimere . . . . .	144
<b>6.</b>	<b>Zusammenfassung</b>	<b>149</b>
6.1.	Diskussion . . . . .	149
6.2.	Ausblick . . . . .	150
<b>A.</b>	<b>Anhang</b>	<b>151</b>
A.1.	Erzeugung der Raumdiskretisierung . . . . .	151
A.1.1.	Optimierungsverfahren . . . . .	153
A.2.	Erzeugung der runden Elemente für die Kugel . . . . .	154
	<b>Literaturverzeichnis</b>	<b>155</b>
	<b>Abbildungsverzeichnis</b>	<b>163</b>

Tabellenverzeichnis	167
Publikationsverzeichnis	169
Danksagung	171



# 1. Einleitung

„Das Staunen ist die Einstellung eines Menschen,  
der die Weisheit wahrhaft liebt,  
ja es gibt keinen anderen Anfang der Philosophie als diesen.“  
– PLATON: „Theaitetos“

Seit Anbeginn der Zeit ist der Mensch von Neugier geprägt. So ist es die Neugier, die immer wieder die Wissbegierigen dazu bringt, Phänomenen auf den Grund zu gehen. Dennoch braucht es Zeit bis, sich aus den Bruchstücken der Erkenntnis eine zusammenhängende Theorie formen lässt. So war es ein langer Weg von den Anfängen bis zur „modernen“ Elektrodynamik im Jahre 1865 [38], auf dem wohl die Wenigsten ahnten, welche Bedeutungen die Optik und die Elektronik für die heutige Menschheit besitzen. Und trotz dessen, dass diese Theorie nun mittlerweile 150 Jahre alt ist, hat sie noch nichts von ihrer Faszination verloren. So zeigen Metallnanopartikel bei Bestrahlung mit weißem Licht unterschiedliche Farben. Ursache sind kollektive Elektronenschwingungen im Metall, die Plasmonen. Von G. MIE [42] wurde erstmals eine vollständige analytische Theorie zur Beschreibung dieser Eigenmoden für sphärische Partikel angegeben. Trotz der Erweiterung dieser Theorie auf ähnliche Systeme, sind analytische Lösungen oft nur für stark vereinfachte Systeme zu gewinnen und bedürfen dennoch eines Computers, um die Streueffizienzen zu berechnen. Obgleich sich so eine „semi“-analytische Lösung schnell berechnen lässt, muss für jedes System ein neues Programm geschrieben werden, sodass recht schnell die Suche nach einem universellen Algorithmus begann. Mit der zeitabhängigen Finiten-Differenzen-Methode (FDTD, engl.: „Finite Difference Time Domain“) [72] schien so ein Verfahren gefunden zu sein. Dieses Verfahren, welches sich speziell in der Mikrowellentheorie bewährte, stößt mit seiner regelmäßigen Diskretisierung schnell an seine Grenzen, wenn Problemstellungen mit komplexen Geometrien behandelt werden sollen. Dazu wurde die Finite-Elemente-Methode (FEM) [21] aus dem Bereich der mechanischen Strukturanalyse für die zeitharmonische Elektrodynamik adaptiert.

Mit den fortschreitenden Erkenntnissen über die linearen Wechselwirkungsphänomene von Licht mit plasmonischen Nanosystemen und der Entwicklung von preiswerten gepulsten Lasern änderte sich die wissenschaftliche Fragestellung hin zu neuen Phänomenen, wie der Frequenzverdopplung (SHG, engl.: „Second Harmonic Generation“). Diese nichtlineare Eigenschaft ist speziell im Bereich der Plasmonik von besonderem Interesse, da Metalle einerseits aufgrund ihrer Zentrosymmetrie keine inhärente SHG zeigen sollten und andererseits, durch die Symmetriebrechung an der Partikeloberfläche und den in der Plasmonik vorkommenden großen Nahfeldüberhöhungen, eine Vervielfachung der SHG zeigen könnten.

Diese Phänomene stellen gleichzeitig neue Anforderungen an die Numerik, welche nun gleichzeitig zeitdynamische Prozesse und komplexe, oberflächensensitive Geometrien behandeln muss. Da lag es nahe, die positiven Eigenschaften der bekannten Methoden (FDTD und FEM) zu kombinieren. Diese zeitabhängige, lokale Finite-Element-Methode, das Unstetige Galerkin-Verfahren, wurde erstmals 1973 zur Lösung der Neutronentransportgleichung vorgestellt [56]. Seitdem wurde es auf die Elektrodynamik übertragen [15], erweitert [17, 20, 31, 35] und auf zahlreiche Systeme der Nanooptik angewendet [47, 63]. Neue Forschungen auf diesem Gebiet beschäftigen sich mit Ansätzen zur Simulation bewegter Ladung in elektromagnetischen Feldern [18, 37, 66] und der Steigerung der Leistungsfähigkeit durch die Verwendung von Graphikprozessoren (GPUs) [26].

Im Rahmen dieser Arbeit wurde das grundlegende Unstetige Galerkin-Verfahren für die MAXWELL- und EULER-Gleichungen implementiert (Kapitel 3). In Kapitel 4 wird die Konstruktion der runden Elemente zur oberflächenkonformen Modellierung nanoplasmonischer Systeme gezeigt. Das letzte Kapitel beschäftigt sich mit den Auswirkungen der runden Elemente und deren Anwendung für die Nanoplasmonik sowie mit den Untersuchungen zur Frequenzverdopplung.

## 2. Grundlagen

Vor der speziellen Diskussion der unstetigen GALERKIN-Methode, gekrümmter Elemente oder kompletten nanooptischer Simulationen sollen in diesem Kapitel die erforderlichen Grundlagen zusammengefasst dargestellt werden. Dabei handelt es sich um grundlegende Differentialgleichungssysteme zur Beschreibung elektromagnetischer Wellen und optischer Materialmodelle. Im zweiten Teil des Kapitels werden die grundlegenden Eigenschaften plasmonischer Systeme in der Nanooptik behandelt, sowie das generelle Vorgehen zur numerischen Berechnung nanooptischer Systeme vorgestellt.

### 2.1. Die MAXWELL - Gleichungen

Der für diese Arbeit wichtigste Gleichungssatz behandelt unter anderem die Ausbreitung elektromagnetischer Wellen. Diese MAXWELL-Gleichungen<sup>1</sup> lauten unter Berücksichtigung von freien Ladungen und Strömen sowie bei Anwesenheit von Materie [4, 38, 49]:

$$\vec{\nabla} \cdot \vec{D}(\vec{r}, t) = \rho_F(\vec{r}, t) \quad (2.1.1a)$$

$$\vec{\nabla} \cdot \vec{B}(\vec{r}, t) = 0 \quad (2.1.1b)$$

$$\vec{\nabla} \times \vec{E}(\vec{r}, t) = -\partial_t \vec{B}(\vec{r}, t) \quad (2.1.1c)$$

$$\vec{\nabla} \times \vec{H}(\vec{r}, t) = \partial_t \vec{D}(\vec{r}, t) + \vec{j}_F(\vec{r}, t) \quad (2.1.1d)$$

In diesem Gleichungssatz sind  $\vec{E}$  die elektrische Feldstärke,  $\vec{B}$  die magnetische Flussdichte,  $\vec{D}$  die elektrische Flussdichte<sup>2</sup> und  $\vec{H}$  die magnetische Feldstärke, sowie  $\rho_F$  und  $\vec{j}_F$  die freie Ladungsdichte bzw. die freie Stromdichte. Diese Gleichungen sind in dieser Form noch unvollständig, da der Zusammenhang von  $\vec{E}$ ,  $\vec{D}$ ,  $\vec{H}$  und  $\vec{B}$  fehlt. Dieser wird erst durch die

---

<sup>1</sup>Benannt nach JAMES CLERK MAXWELL (\*1831 – †1879), der diese Gleichungen zwischen 1861-1864 zusammenstellte und in [38] veröffentlichte.

<sup>2</sup>Die elektrische Flussdichte wird häufig auch dielektrische Verschiebung genannt, so dass  $\partial_t \vec{D}$  als Verschiebungsstromdichte bezeichnet wird.

Materialgleichungen:

$$\vec{D}(\vec{r}, t) = \epsilon_0 \vec{E}(\vec{r}, t) + \vec{P}(\vec{r}, t) \quad (2.1.2a)$$

$$\vec{H}(\vec{r}, t) = \frac{1}{\mu_0} \vec{B}(\vec{r}, t) - \vec{M}(\vec{r}, t) \quad (2.1.2b)$$

hergestellt. Dabei wurden die Polarisation  $\vec{P}$ , die Magnetisierung  $\vec{M}$  und die (einheitenssystemabhängigen) Naturkonstanten Vakuumpermeabilität  $\mu_0$  beziehungsweise Vakuumpermittivität  $\epsilon_0$  eingeführt. Da bei den in dieser Arbeit verwendeten Systemen die Magnetisierung eine untergeordnete Rolle spielt, wird diese hier nicht weiter betrachtet. Nichtsdestotrotz lassen sich alle im weiteren Verlauf für die Polarisation hergeleiteten Zusammenhänge sinngemäß auch auf die Magnetisierung anwenden, so dass speziell für die Permeabilität eine ähnliche Notation wie für die Permittivität verwendet wird, ohne jedoch die komplette Herleitung mit anzugeben.

Für den weiteren Verlauf ist es hilfreich, die Polarisation  $\vec{P}$  noch differenzierter zu betrachten:

$$\vec{P}(\vec{r}, t) = \epsilon_0 \chi_E^\infty(\vec{r}) \vec{E}(\vec{r}, t) + \underbrace{\sum_i \vec{P}_i^{(L)}(\vec{r}, t) + \sum_i \vec{P}_i^{(NL)}(\vec{r}, t)}_{\sum_i \vec{P}_i(\vec{r}, t)} \quad (2.1.3)$$

Dabei beschreibt  $\chi_E$  die elektrische Suszeptibilität<sup>3</sup>. Diese im Allgemeinen nichtlineare, tensorielle, frequenz- und zeitabhängige Verknüpfung des elektrischen Feldes mit der Polarisation (siehe dazu Kapitel 2.1.2 und 2.5) wird hier ( $\chi_E \rightarrow \chi_E^\infty$ ) als frequenzunabhängig, isotrop und linear betrachtet und dient der Beschreibung einer konstanten Hintergrundpolarisierbarkeit. Die weiteren Terme  $\vec{P}_i^{(L)}$  und  $\vec{P}_i^{(NL)}$  beschreiben die zeitabhängige, tensorielle, lineare und nichtlineare Polarisation. Mit den Abkürzungen:

$$\epsilon^\infty = \epsilon_0 \epsilon_r^\infty = \epsilon_0 (1 + \chi_E^\infty) \quad \mu^\infty = \mu_0 \mu_r^\infty = \mu_0 (1 + \chi_M^\infty) \quad (2.1.4)$$

und der Vernachlässigung der freien Ladungs- und Stromdichten ( $\rho_F = 0$  und  $\vec{j}_F = 0$ ) und mit  $\vec{j}$  als Polarisationsstromdichte und der linearen magnetischen Antwort  $\vec{M}(\vec{r}, t) =$

---

<sup>3</sup>Das magnetische Pendant  $\chi_M^\infty$  heißt dementsprechend magnetische Suszeptibilität.



$\chi_M^\infty(\vec{r})\vec{H}(\vec{r}, t)$  lauten die MAXWELL-Gleichungen nun:

$$\vec{\nabla} \cdot \epsilon^\infty \vec{E}(\vec{r}, t) = - \sum_i \vec{\nabla} \cdot \vec{P}_i(\vec{r}, t) \quad (2.1.5a)$$

$$\vec{\nabla} \cdot \mu^\infty \vec{H}(\vec{r}, t) = 0 \quad (2.1.5b)$$

$$\vec{\nabla} \times \vec{E}(\vec{r}, t) = -\mu^\infty \partial_t \vec{H}(\vec{r}, t) \quad (2.1.5c)$$

$$\vec{\nabla} \times \vec{H}(\vec{r}, t) = \epsilon^\infty \partial_t \vec{E}(\vec{r}, t) + \underbrace{\sum_i \partial_t \vec{P}_i(\vec{r}, t)}_{\vec{j}(\vec{r}, t)} \quad (2.1.5d)$$

### 2.1.1. MAXWELL-Gleichung in niedrigeren Dimensionen

Häufig findet man den Ausdruck ein- bzw. zweidimensionale MAXWELL-Gleichungen. Dies ist insofern verwirrend, da es die MAXWELL-Theorie zur Beschreibung der Elektrodynamik nur in drei Dimensionen gibt. Wenn jedoch bestimmte Symmetriebedingungen erfüllt sind, lassen sich diese Gleichungen in eine einfachere Form bringen, was im Folgenden gezeigt werden soll.

#### Die MAXWELL-Gleichungen in 2D

Setzt man eine Translationssymmetrie in  $z$ -Richtung voraus und nimmt an, dass sich die Felder nur in der  $xy$ -Ebene ausbreiten, dann sind die Felder in  $z$ -Richtung von der Raumkoordinate unabhängig. Dies entspricht der Modellierung von Strukturen, die in einer Dimension sehr weit („unendlich“) ausgedehnt sind. Damit verschwindet die  $z$ -Ableitung aus dem Nabla-Operator und die inhomogenen MAXWELL-Gleichungen (2.1.5c,d) zerfallen in zwei voneinander unabhängige Gleichungssätze, erstens der  $TE$ -Fall<sup>4</sup>:

$$\mu^\infty \partial_t H_z = \partial_y E_x - \partial_x E_y \quad (2.1.6a)$$

$$\epsilon^\infty \partial_t E_x = +\partial_y H_z - j_x \quad (2.1.6b)$$

$$\epsilon^\infty \partial_t E_y = -\partial_x H_z - j_y \quad (2.1.6c)$$

---

<sup>4</sup>Benannt nach den nun transversalen  $\vec{E}_\perp = E_{x,y}$ -Feldkomponenten

und zweitens, der *TM*-Fall:

$$\epsilon^\infty \partial_t E_z = \partial_x H_y - \partial_y H_x - j_z \quad (2.1.7a)$$

$$\mu^\infty \partial_t H_x = -\partial_y E_z \quad (2.1.7b)$$

$$\mu^\infty \partial_t H_y = +\partial_x E_z \quad (2.1.7c)$$

### Die MAXWELL-Gleichungen in 1D

Wenn nun neben der Translationssymmetrie entlang der  $z$ -Achse, auch noch die entlang der  $y$ -Achse und dementsprechend eine Ausbreitung in  $x$ -Richtung vorliegt, gelangt man zur eindimensionalen Form:

$$\epsilon^\infty \partial_t E_y = -\partial_x H_z - j_y \quad (2.1.8a)$$

$$\mu^\infty \partial_t H_z = -\partial_x E_y \quad (2.1.8b)$$

Dabei genügt hier die Betrachtung von nur noch zwei der ursprünglich sechs Feldkomponenten.

#### 2.1.2. Die zeitharmonischen MAXWELL-Gleichungen

Für ein elektromagnetisches System im eingeschwungenen Zustand, das heißt, nach einer hinreichend langen Einschwingzeit, kann man annehmen, dass es mit der Frequenz  $\omega$  oszilliert. Damit lässt sich jede Feldvariable durch  $\vec{\mathcal{F}}(\vec{r}, t) = \Re(\check{\vec{\mathcal{F}}}_0(\vec{r}) e^{-i\omega t})$  darstellen<sup>5</sup>. Das Brevis ( $\check{\phantom{x}}$ ) bezeichnet hier die komplexen Zahlen und Vektoren. Mit diesem Ansatz lauten die MAXWELL-Gleichungen ohne freie Ladungen und Ströme:

$$\vec{\nabla} \cdot \check{\vec{D}}(\vec{r}, \omega) = 0 \quad (2.1.9a)$$

$$\vec{\nabla} \cdot \check{\vec{B}}(\vec{r}, \omega) = 0 \quad (2.1.9b)$$

$$\vec{\nabla} \times \check{\vec{E}}(\vec{r}, \omega) = i\omega \check{\vec{B}}(\vec{r}, \omega) \quad (2.1.9c)$$

$$\vec{\nabla} \times \check{\vec{H}}(\vec{r}, \omega) = -i\omega \check{\vec{D}}(\vec{r}, \omega) \quad (2.1.9d)$$

<sup>5</sup>Die konkrete Wahl der Zeitabseparation  $\check{\vec{\mathcal{F}}}(\vec{r}, t) = \check{\vec{\mathcal{F}}}_0(\vec{r}) e^{-i\omega t}$  bzw.  $\check{\vec{\mathcal{F}}}(\vec{r}, t) = \check{\vec{\mathcal{F}}}_0(\vec{r}) e^{+i\omega t}$  ist prinzipiell willkürlich. Hat man sich jedoch für eine Variante entschieden, muss diese für alle Felder einheitlich sein, da ein Übergang von  $e^{-i\omega t} \rightarrow e^{+i\omega t}$  einer unphysikalischen Zeitumkehr entspricht.

Um nun die zugehörigen Materialgleichungen herzuleiten, werden nur homogene, isotrope, lineare, lokale Materialien berücksichtigt<sup>6</sup>. Ein allgemeinerer Fall wird in Abschnitt 2.5 behandelt. Mit  $\vec{E}(\vec{r}, t) = \int d\omega \check{\vec{E}}(\vec{r}, \omega) e^{-i\omega t}$  transformiert sich die lineare Polarisation  $P^{(L)}$  zu [43]:

$$\begin{aligned}
 \vec{P}(\vec{r}, t) &= \epsilon_0 \int dt' \chi_E(\vec{r}, t - t') \vec{E}(\vec{r}, t') \\
 &= \epsilon_0 \int d\omega \int dt' \chi_E(\vec{r}, t - t') \check{\vec{E}}_0(\vec{r}, \omega) e^{-i\omega t'} \\
 &= \epsilon_0 \int d\omega \underbrace{\int dt'' \chi_E(\vec{r}, t'') e^{i\omega t''}}_{\check{P}(\vec{r}, \omega) = \check{\chi}_E(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega)} \check{\vec{E}}_0(\vec{r}, \omega) e^{-i\omega t}
 \end{aligned} \tag{2.1.10}$$

Damit folgt für die zeitharmonische Polarisation<sup>7</sup>:

$$\begin{aligned}
 \check{\vec{P}}(\vec{r}, \omega) &= \epsilon_0 \chi_E^\infty(\vec{r}) \check{\vec{E}}(\vec{r}, \omega) + \sum_i \check{\vec{P}}_i^{(L)}(\vec{r}, \omega) \\
 &= \epsilon_0 \chi_E^\infty(\vec{r}) \check{\vec{E}}(\vec{r}, \omega) + \sum_i \epsilon_0 \check{\chi}_{E,i}(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega) \\
 &= \epsilon_0 \left( \chi_E^\infty(\vec{r}) + \sum_i \check{\chi}_{E,i}(\vec{r}, \omega) \right) \check{\vec{E}}(\vec{r}, \omega) \\
 &= \epsilon_0 \check{\chi}_E(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega)
 \end{aligned} \tag{2.1.11}$$

---

<sup>6</sup>Lineare Materialien lassen sich durch

$$\vec{P}(\vec{r}, t) = \epsilon_0 \int d^3\vec{r}' \int_{-\infty}^t dt' \overset{\leftrightarrow}{\chi}_E(\vec{r} - \vec{r}', t - t') \vec{E}(\vec{r}', t')$$

ausdrücken. Hier ist zu beachten, dass die Suszeptibilität nicht von  $\vec{E}$  oder  $\vec{H}$  abhängt. Ist das Material isotrop, so reduziert sich der Tensor  $\overset{\leftrightarrow}{\chi}_E$  zu einer skalaren Größe  $\chi_E$ . Für eine lokale Materialeigenschaft gilt zusätzlich noch  $\chi_E(\vec{r} - \vec{r}', t - t') = \chi_E(\vec{r}, t - t') \delta(\vec{r} - \vec{r}')$ , das heißt:

$$\vec{P}(\vec{r}, t) = \epsilon_0 \int_{-\infty}^t dt' \chi_E(\vec{r}, t - t') \vec{E}(\vec{r}, t')$$

<sup>7</sup>Für die Magnetisierung ergibt sich sinngemäß:  $\check{\vec{M}}(\vec{r}, \omega) = \check{\chi}_M(\vec{r}, \omega) \check{\vec{H}}(\vec{r}, \omega)$

Damit lassen sich die Materialgleichungen weiter vereinfachen zu:

$$\check{\vec{D}}(\vec{r}, \omega) = \epsilon_0 \check{\vec{E}}(\vec{r}, \omega) + \check{\vec{P}}(\vec{r}, \omega) \quad (2.1.12a)$$

$$= \epsilon_0 (1 + \check{\chi}_E(\vec{r}, \omega)) \check{\vec{E}}(\vec{r}, \omega) = \epsilon_0 \check{\epsilon}_r(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega) = \check{\epsilon}(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega)$$

$$\check{\vec{B}}(\vec{r}, \omega) = \mu_0 (\check{\vec{H}}(\vec{r}, \omega) + \check{\vec{M}}(\vec{r}, \omega)) \quad (2.1.12b)$$

$$= \mu_0 (1 + \check{\chi}_M(\vec{r}, \omega)) \check{\vec{H}}(\vec{r}, \omega) = \mu_0 \check{\mu}_r(\vec{r}, \omega) \check{\vec{H}}(\vec{r}, \omega) = \check{\mu}(\vec{r}, \omega) \check{\vec{H}}(\vec{r}, \omega)$$

Setzt man diese jetzt direkt in die Gleichung 2.1.9 ein, erhält man:

$$\vec{\nabla} \cdot (\check{\epsilon}(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega)) = 0 \quad (2.1.13a)$$

$$\vec{\nabla} \cdot (\check{\mu}(\vec{r}, \omega) \check{\vec{H}}(\vec{r}, \omega)) = 0 \quad (2.1.13b)$$

$$\vec{\nabla} \times \check{\vec{E}}(\vec{r}, \omega) = i\omega \check{\mu}(\vec{r}, \omega) \check{\vec{H}}(\vec{r}, \omega) \quad (2.1.13c)$$

$$\vec{\nabla} \times \check{\vec{H}}(\vec{r}, \omega) = -i\omega \check{\epsilon}(\vec{r}, \omega) \check{\vec{E}}(\vec{r}, \omega) \quad (2.1.13d)$$

Eine Aufgabe von Abschnitt 2.3 wird es sein, geeignete Materialmodelle in Form von  $\check{\vec{P}}(\vec{r}, t)$  bzw.  $\check{\epsilon}_r(\vec{r}, \omega)$  bereitzustellen, um die MAXWELL-Gleichung bei dispersiven Materialien wie z.B. Metallen behandeln zu können.

### 2.1.3. Stetigkeit der Felder

Da eine realistische Simulation in aller Regel Materialgrenzflächen benötigt, müssen geeignete Übergangsbedingungen für die Felder abgeleitet werden. Dabei macht man sich die Stetigkeit einzelner Feldkomponenten als Folge des STOKESSchem und des GAUSSSchem Satzes beim Materialübergang ( $1 \leftrightarrow 2$ ) zu nutze. Mit  $\vec{n}$  als Normalenvektor der Materialgrenze ergibt sich aus der Divergenz (unter der Voraussetzung verschwindender Oberflächenladungsdichte und mit  $\vec{D} = \epsilon \vec{E}$  und  $\vec{B} = \mu \vec{H}$ ):

$$\vec{n} \cdot \epsilon_2 \vec{E}_2 = \vec{n} \cdot \epsilon_1 \vec{E}_1 \quad (2.1.14a)$$

$$\vec{n} \cdot \mu_2 \vec{H}_2 = \vec{n} \cdot \mu_1 \vec{H}_1 \quad (2.1.14b)$$

und aus der Rotation (bei Verschwinden der Oberflächenstromdichte):

$$\vec{n} \times \vec{E}_2 = \vec{n} \times \vec{E}_1 \quad (2.1.15a)$$

$$\vec{n} \times \vec{H}_2 = \vec{n} \times \vec{H}_1 \quad (2.1.15b)$$

Diese Stetigkeitsbedingungen gelten gleichermaßen für die Zeit- und die Frequenzdomäne.

### 2.1.4. Lösungen der MAXWELL-Gleichungen

Neben den allgemeinen Betrachtungen zu den MAXWELL-Gleichungen sind auch spezielle, analytische Lösungen für diese Arbeit von Bedeutung, einerseits, um die notwendigen Randbedingungen der Simulationen zu setzen und andererseits als Referenz für die numerischen Ergebnisse. Im Folgenden werden daher ausgewählte Grundlösungen der MAXWELL-Gleichungen vorgestellt.

#### Vorbetrachtung

Da alle hier behandelten Lösungen auf der Wellengleichung des elektromagnetischen Feldes beruhen, wird diese zunächst hergeleitet. Dabei werden ausschließlich lineare, homogene und isotrope Materialkonstanten  $\epsilon$  und  $\mu$  berücksichtigt. Ein gängiger Ansatz (siehe [4, 45, 49]) ist es Gleichung 2.1.13c in Gleichung 2.1.13d einzusetzen und die Regel für das doppelte Kreuzprodukt<sup>8</sup> sowie die Divergenzfreiheit des Feldes auszunutzen<sup>9</sup>:

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = i\omega\mu\vec{\nabla} \times \vec{H} = \omega^2\epsilon\mu\vec{E} \quad (2.1.16a)$$

$$\vec{\nabla}^2 \vec{E} + \omega^2\epsilon\mu\vec{E} = 0 \quad (2.1.16b)$$

Die letzte Gleichung heißt auch HELMHOLTZ-Gleichung<sup>10</sup>. Als nächstes führt man den Wellenvektor  $\vec{k}$  oder die Wellenzahl  $k$  ein<sup>11</sup>:

$$\vec{k}^2 = \omega^2\epsilon\mu \quad k^2 = \omega^2\epsilon\mu \quad (2.1.17)$$

Welche der beiden Möglichkeiten gewählt wird, hängt von der konkreten Fragestellung ab. Da jedoch  $\vec{k}^2 = k^2$  ist, gelten Aussagen bezüglich  $\vec{k}^2$  auch immer für  $k^2$ . Mit Hilfe der Lichtgeschwindigkeit im Medium  $c$ , beziehungsweise der Lichtgeschwindigkeit im Vakuum

---

<sup>8</sup> $A \times (B \times C) = B(A \cdot C) - C \cdot (AB)$

<sup>9</sup>Für das  $H$ -Feld lauten die Gleichungen:

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{H}) = -i\omega\epsilon\vec{\nabla} \times \vec{E} = \omega^2\epsilon\mu\vec{H}$$

$$\vec{\nabla}^2 \vec{H} + \omega^2\epsilon\mu\vec{H} = 0$$

<sup>10</sup>Nach HERMANN LUDWIG FERDINAND VON HELMHOLTZ (\* 31. August 1821 in Potsdam; † 8. September 1894 in Charlottenburg)

<sup>11</sup>Hier muss beachtet werden, dass  $\vec{k} \cdot \vec{k}$  kein Skalarprodukt ist, da  $\mathbb{C} \ni \omega^2\epsilon\mu = \vec{k} \cdot \vec{k} \neq \langle \vec{k}, \vec{k} \rangle \in \mathbb{R}$ . Insbesondere ist  $\mathbb{R} \ni |\vec{k}| \neq k \in \mathbb{C}$ .

$c_0$  und dem Brechungsindex  $\check{n}$  folgt:

$$\frac{1}{\check{c}} = \sqrt{\check{\epsilon}\check{\mu}} = \sqrt{\check{\epsilon}_r\check{\mu}_r}\sqrt{\epsilon_0\mu_0} = \check{n}\frac{1}{c_0} \quad (2.1.18)$$

Damit ergibt sich ein Zusammenhang zwischen der Lichtgeschwindigkeit, den Materialkonstanten und der Wellenlänge  $\check{\lambda}$ :

$$\check{k}^2 = \frac{\omega^2}{c^2} \quad \check{\lambda}^2 = \frac{(2\pi)^2}{\check{k}^2} \quad (2.1.19)$$

Da sich die Größen Wellenlänge, Frequenz und Photonenenergie leicht ineinander umrechnen lassen, wird die Frequenz  $\omega$  häufig in Einheiten der Vakuumwellenlänge ( $\lambda_0 = 2\pi c_0/\omega$ ) oder der Energie ( $\hbar\omega$ ) statt [1/s] angegeben.

### Kartesische Koordinaten - Ebene Welle

Die bekannteste Lösung der MAXWELL-Gleichungen ist die ebene Welle. Der Ansatz:

$$\check{\vec{E}} = \check{\vec{E}}_0 e^{i(\check{k}\vec{r} - \omega t)} \quad (2.1.20)$$

löst obige Wellengleichung. Dabei ist  $\check{k}$  der komplexe Wellenvektor mit  $\check{k}^2 = \omega^2\check{\epsilon}\check{\mu}$ . Der Imaginärteil  $\Im(\check{k})$  beschreibt hierbei die Richtung und Stärke der Absorption und der Realteil  $\Re(\check{k})$  die Richtung und Wellenlänge der Propagation. Durch Einsetzen der Lösung in die Gleichungen 2.1.13 erhält man:

$$\check{k} \times \check{\vec{H}}_0 = -\omega\check{\epsilon}\check{\vec{E}}_0 \quad (2.1.21a)$$

$$\check{k} \times \check{\vec{E}}_0 = \omega\check{\mu}\check{\vec{H}}_0 \quad (2.1.21b)$$

$$\check{k} \cdot \check{\vec{E}}_0 = \check{k} \cdot \check{\vec{H}}_0 = \check{\vec{E}}_0 \cdot \check{\vec{H}}_0 = 0 \quad (2.1.21c)$$

Das heißt, dass  $\check{\vec{E}}_0$ ,  $\check{\vec{H}}_0$  und  $\check{k}$ , für den Fall, dass sie reell sind, senkrecht aufeinander stehen<sup>12</sup>. Für den Fall, dass sich der Wellenvektor (mit  $\vec{e}_{\check{k}}$  als reellem Einheitsvektor in Propagationsrichtung) als  $\check{k} = \omega\sqrt{\check{\epsilon}\check{\mu}}\vec{e}_{\check{k}}$  schreiben lässt, heißt die Lösung ebene Welle. Für

---

<sup>12</sup>Die Beziehung  $\check{k} \cdot \check{\vec{E}}_0 = \check{k} \cdot \check{\vec{H}}_0 = \check{\vec{E}}_0 \cdot \check{\vec{H}}_0 = 0$  gilt zwar allgemein, allerdings ist der Begriff „senkrecht“ im komplexen nicht so intuitiv.

diese ist nach Gleichung 2.1.21 der Zusammenhang zwischen  $\check{\vec{E}}_0$  und  $\check{\vec{H}}_0$  gegeben durch:

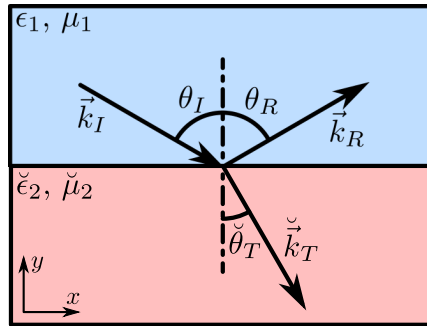
$$\vec{e}_k \times \check{\vec{H}}_0 = \sqrt{\frac{\check{\epsilon}}{\check{\mu}}} \check{\vec{E}}_0 = \check{Y} \check{\vec{E}}_0 = \frac{1}{\check{Z}} \check{\vec{E}}_0 \quad (2.1.22)$$

Die neu eingeführten Werte  $\check{Z} = \sqrt{\check{\mu}/\check{\epsilon}}$  und  $\check{Y} = \sqrt{\check{\epsilon}/\check{\mu}}$  werden als Impedanz und Admittanz bezeichnet und geben den quantitativen Zusammenhang zwischen dem  $E$ - und dem  $H$ -Feld an.

Aus dieser Grundlösung lassen sich, speziell wenn  $\check{k} \neq \omega \sqrt{\check{\epsilon}\check{\mu}} \vec{e}_k$  ist, weitere Wellenformen ableiten, welche insbesondere bei Wellen an Materialgrenzflächen eine bedeutende Rolle spielen. Eine Klasse davon sind die evaneszenten Wellen<sup>13</sup>, welche in Abschnitt 2.4.1 benötigt werden.

### Brechung und Reflexion an einer Grenzfläche

Mit Hilfe der ebenen Wellen, lassen sich nun Lösungen für spezielle Geometrien herleiten. Eine der wichtigsten ist die Brechung und Reflexion einer ebenen Welle an einer ebenen Grenzfläche (siehe Abbildung 2.1.1). Dabei genügt es, hier davon auszugehen, dass die Domäne 1 durch reelle Materialparameter beschrieben werden kann<sup>14</sup>. Als Ansatz dienen



**Abbildung 2.1.1.:** Brechung und Reflexion an einer ebenen Grenzfläche, mit der einfallenden Welle  $I$ , der reflektierten Welle  $R$  und der transmittierten Welle  $T$ .

hier drei Wellen, eine einfallende ( $I$ ), eine reflektierte ( $R$ ) und eine transmittierte Welle ( $T$ ). Aus Gründen der Symmetrie genügt es, anzunehmen, dass der Wellenvektor der einfallenden Welle die Form  $\vec{k}_I = k_x \vec{e}_x + k_y \vec{e}_y$  besitzt. Das Koordinatensystem wird so

<sup>13</sup>Bei diesen Wellen ist  $\Re(\check{k}) \cdot \Im(\check{k}) = 0$ . Das heißt, der imaginäre und der reelle Anteil des Wellenvektors stehen senkrecht aufeinander.

<sup>14</sup>Die Herleitung ist in ähnlicher Form auch für komplexe  $\check{\epsilon}_1$  und  $\check{\mu}_1$  gültig. Da Medium 1 typischerweise ein nicht- beziehungsweise schwachabsorbierendes Medium ist, wird hier nur der Spezialfall behandelt.

gewählt, dass die Grenzfläche durch den Ursprung geht und die Grenzflächennormale in  $+y$ -Richtung zeigt ( $\vec{n} = \vec{e}_y$ ). Aus der Beibehaltung der Frequenz folgt:

$$\frac{\vec{k}_I^2}{\epsilon_1 \mu_1} = \frac{\vec{k}_R^2}{\epsilon_1 \mu_1} = \frac{\check{\vec{k}}_T^2}{\check{\epsilon}_2 \check{\mu}_2} = \omega^2 \quad (2.1.23)$$

Da die Stetigkeitsbedingung (Abschnitt 2.1.3) zu jeder Zeit erfüllt sein muss, gilt:

$$\vec{k}_I \cdot \vec{r} = \vec{k}_R \cdot \vec{r} = \check{\vec{k}}_T \cdot \vec{r} = \vec{k}_\parallel \cdot \vec{r} \quad \forall \vec{r} \text{ mit } r_y = 0 \quad (2.1.24)$$

Das heißt, dass der Wellenvektoranteil parallel zur Grenzfläche  $\vec{k}_\parallel$  mit dem senkrechten Anteil des Wellenvektors der einfallenden Welle  $\vec{k}_{\perp,I} = -(\vec{k}_I^2 - \vec{k}_\parallel^2)^{1/2} \vec{e}_y = k_{y,I} \vec{e}_y$  eine Ebene, die Einfallsebene, bildet. Des Weiteren folgt aus den beiden Gleichungen, dass  $\vec{k}_{\perp,1} = \vec{k}_{\perp,R} = -\vec{k}_{\perp,I}$  und  $\check{\vec{k}}_{\perp,2} = \check{\vec{k}}_{\perp,T} = -(\check{\vec{k}}_T^2 - \vec{k}_\parallel^2)^{1/2} \vec{e}_y$  sein muss<sup>15</sup>.

Nun bietet es sich an, zwischen zwei Polarisationsarten zu unterscheiden: die s-Polarisation mit dem  $E$ -Feld senkrecht zur Einfallsebene, und die p-Polarisation mit dem  $E$ -Feld parallel zur Einfallsebene. Begonnen wird hier mit der s-Polarisation. Mit den Erkenntnissen aus dem letzten Abschnitt und dem Reflexionskoeffizienten  $\check{r}_s$  und Transmissionskoeffizienten  $\check{t}_s$  lauten die Ansatzfunktionen:

$$\check{\vec{E}}_I = E_0 e^{i(\vec{k}_I \cdot \vec{r} - \omega t)} \vec{e}_z \quad \check{\vec{H}}_I = \frac{\vec{k}_I \times \check{\vec{E}}_I}{\omega \mu_1} = \frac{E_0}{\omega \mu_1} (k_{\perp,1} \vec{e}_x - k_\parallel \vec{e}_y) e^{i(\vec{k}_I \cdot \vec{r} - \omega t)} \quad (2.1.25a)$$

$$\check{\vec{E}}_R = E_0 \check{r}_s e^{i(\vec{k}_R \cdot \vec{r} - \omega t)} \vec{e}_z \quad \check{\vec{H}}_R = \frac{\vec{k}_R \times \check{\vec{E}}_R}{\omega \mu_1} = \frac{-E_0 \check{r}_s}{\omega \mu_1} (k_{\perp,1} \vec{e}_x + k_\parallel \vec{e}_y) e^{i(\vec{k}_R \cdot \vec{r} - \omega t)} \quad (2.1.25b)$$

$$\check{\vec{E}}_T = E_0 \check{t}_s e^{i(\check{\vec{k}}_T \cdot \vec{r} - \omega t)} \vec{e}_z \quad \check{\vec{H}}_T = \frac{\check{\vec{k}}_T \times \check{\vec{E}}_T}{\omega \check{\mu}_2} = \frac{E_0 \check{t}_s}{\omega \check{\mu}_2} (\check{k}_{\perp,2} \vec{e}_x - k_\parallel \vec{e}_y) e^{i(\check{\vec{k}}_T \cdot \vec{r} - \omega t)} \quad (2.1.25c)$$

<sup>15</sup>Damit folgt auch das Reflexionsgesetz:

$$\theta_I = \theta_R$$

und Brechungsgesetz nach SNELLIUS:

$$\sqrt{\epsilon_1 \mu_1} \sin(\theta_I) = \sqrt{\check{\epsilon}_2 \check{\mu}_2} \sin(\check{\theta}_T)$$

Dabei ist der Transmissionswinkel im Allgemeinen komplex.



Mit Hilfe der Stetigkeitsbedingung (Abschnitt 2.1.3) folgt:

$$1 + \check{r}_s = \check{t}_s \quad \text{Stetigkeit von E-Feld } \parallel \text{ bzw. H-Feld } \perp \quad (2.1.26a)$$

$$\frac{1}{\mu_1} k_{\perp,1} (1 - \check{r}_s) = \frac{1}{\check{\mu}_2} \check{k}_{\perp,2} \check{t}_s \quad \text{Stetigkeit von H-Feld } \parallel \quad (2.1.26b)$$

Mit dem Einfallswinkel  $\tan(\theta_I) = k_{\parallel}/k_{\perp}$  und dem komplexen Brechungsindex  $\check{n} = \sqrt{\check{\epsilon}_r \check{\mu}_r}$  erhält man schließlich die FRESNELSchen-Formeln [49]:

$$\check{r}_s = \frac{n_1 \cos(\theta_I) - \frac{\mu_1}{\check{\mu}_2} \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}}{n_1 \cos(\theta_I) + \frac{\mu_1}{\check{\mu}_2} \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}} \quad (2.1.27a)$$

$$\check{t}_s = \frac{2n_1 \cos(\theta_I)}{n_1 \cos(\theta_I) + \frac{\mu_1}{\check{\mu}_2} \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}} \quad (2.1.27b)$$

Für die p-Polarisation kommt man analog auf [49]:

$$\check{r}_p = \frac{\frac{\mu_1}{\check{\mu}_2} \check{n}_2^2 \cos(\theta_I) - n_1 \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}}{\frac{\mu_1}{\check{\mu}_2} \check{n}_2^2 \cos(\theta_I) + n_1 \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}} \quad (2.1.28a)$$

$$\check{t}_p = \frac{2n_1 \check{n}_2 \cos(\theta_I)}{\frac{\mu_1}{\check{\mu}_2} \check{n}_2^2 \cos(\theta_I) + n_1 \sqrt{\check{n}_2^2 - n_1^2 \sin^2(\theta_I)}} \quad (2.1.28b)$$

## Eigenlösung Grenzfläche

Neben der Lösung an der Grenzfläche mit Anregung durch eine einfallende Welle soll es nun um die Eigenlösung der Grenzfläche gehen. Dabei wird angenommen, dass beide Medien mit Hilfe komplexer Materialparameter beschrieben werden können. Als Ansatz dienen zwei Wellen, je eine auf jeder Seite der Grenzfläche. Dabei wird mit dem Fall der p-Polarisation begonnen. Die Ansatzfunktionen lauten:

$$\check{\vec{H}}_1 = e^{i(\check{k}_1 \cdot \vec{r} - \omega t)} \vec{e}_z \quad \check{\vec{E}}_1 = \frac{\check{k}_1 \times \check{\vec{H}}_1}{-\omega \check{\epsilon}_1} = \frac{-1}{\omega \check{\epsilon}_1} (\check{k}_{\perp,1} \vec{e}_x - \check{k}_{\parallel} \vec{e}_y) e^{i(\check{k}_1 \cdot \vec{r} - \omega t)} \quad (2.1.29a)$$

$$\check{\vec{H}}_2 = e^{i(\check{k}_2 \cdot \vec{r} - \omega t)} \vec{e}_z \quad \check{\vec{E}}_2 = \frac{\check{k}_2 \times \check{\vec{H}}_2}{-\omega \check{\epsilon}_2} = \frac{-1}{\omega \check{\epsilon}_2} (\check{k}_{\perp,2} \vec{e}_x - \check{k}_{\parallel} \vec{e}_y) e^{i(\check{k}_2 \cdot \vec{r} - \omega t)} \quad (2.1.29b)$$

Dabei wurden bereits die Symmetriebedingung für  $\check{k}_{\parallel} = \check{k}_{1,x} = \check{k}_{2,x}$  und die Stetigkeitsbedingung für  $\check{\vec{H}}_1(\vec{r})|_{r_y=0} = \check{\vec{H}}_2(\vec{r})|_{r_y=0}$  berücksichtigt. Aus der Stetigkeitsbedingung für das

$E$ -Feld folgt<sup>16</sup>:

$$\frac{1}{\check{\epsilon}_1} \check{k}_{\perp,1} = \frac{1}{\check{\epsilon}_2} \check{k}_{\perp,2} \quad (2.1.30a)$$

$$\check{\epsilon}_2 \sqrt{\omega^2 \check{\epsilon}_1 \check{\mu}_1 - \check{k}_{\parallel}^2} = \check{\epsilon}_1 \sqrt{\omega^2 \check{\epsilon}_2 \check{\mu}_2 - \check{k}_{\parallel}^2} \quad (2.1.30b)$$

$$\check{k}_{\parallel}^2 = \frac{\check{\epsilon}_1 \check{\epsilon}_2 \omega^2 (\check{\epsilon}_1 \check{\mu}_2 - \check{\epsilon}_2 \check{\mu}_1)}{(\check{\epsilon}_1 + \check{\epsilon}_2)(\check{\epsilon}_1 - \check{\epsilon}_2)} \quad (2.1.30c)$$

Diese Lösung existiert offenbar nur, wenn es einen Materialkontrast in der Permittivität gibt. In Abschnitt 2.4.1 wird dann aus dieser Lösung das Oberflächenplasmon abgeleitet.

### Krummlinige Koordinaten

Neben den kartesischen Koordinaten werden insbesondere bei der Behandlung von Zylinderbeziehungsweise sphärischen Symmetrien die Lösung der MAXWELL-Gleichungen in krummlinigen Koordinaten benötigt. Die hier vorgestellte Herleitung basiert auf [4]. Als Ausgangspunkte dienen ebenfalls wieder die HELMHOLTZ-Gleichung<sup>17</sup> und die MAXWELL-Gleichungen:

$$\vec{\nabla}^2 \check{\vec{E}} + \check{k}^2 \check{\vec{E}} = 0 \quad \vec{\nabla}^2 \check{\vec{H}} + \check{k}^2 \check{\vec{H}} = 0 \quad (2.1.31a)$$

$$\vec{\nabla} \check{\vec{E}} = 0 \quad \vec{\nabla} \check{\vec{H}} = 0 \quad (2.1.31b)$$

$$\vec{\nabla} \times \check{\vec{E}} = i\omega \check{\mu} \check{\vec{H}} \quad \vec{\nabla} \times \check{\vec{H}} = -i\omega \check{\epsilon} \check{\vec{E}} \quad (2.1.31c)$$

Mit Hilfe einer Skalarfunktion  $\check{\psi}$  und einem Vektor  $\check{\vec{c}}$  konstruiert man sich zunächst eine neue Vektorfunktion  $\check{\vec{\mathcal{F}}}_1$ , so dass:

$$\check{\vec{\mathcal{F}}}_1 = \vec{\nabla} \times (\check{\vec{c}} \check{\psi}) \quad (2.1.32)$$

gilt. Diese erfüllt nun:

$$\vec{\nabla} \check{\vec{\mathcal{F}}}_1 = 0 \quad (2.1.33)$$

---

<sup>16</sup>Die Herleitung für die s-Polarisation liefert:

$$\check{k}_{\parallel}^2 = \frac{\check{\mu}_1 \check{\mu}_2 \omega^2 (\check{\mu}_1 \check{\epsilon}_2 - \check{\mu}_2 \check{\epsilon}_1)}{(\check{\mu}_1 + \check{\mu}_2)(\check{\mu}_1 - \check{\mu}_2)}$$

Diese Lösung existiert nur, wenn  $\check{\mu}_1 \neq \check{\mu}_2$  und ist daher speziell in der Optik von untergeordneter Bedeutung.

<sup>17</sup>Dabei ist  $\check{k}$  nun die Wellenzahl.

Damit erhält man:

$$\vec{\nabla}^2 \check{\vec{\mathcal{F}}}_1 + \check{k}^2 \check{\vec{\mathcal{F}}}_1 = \vec{\nabla} \times [\vec{c}(\vec{\nabla}^2 \check{\psi} + \check{k}^2 \check{\psi})] \quad (2.1.34)$$

Offensichtlich erfüllt  $\check{\vec{\mathcal{F}}}_1$  die HELMHOLTZ-Gleichung genau dann, wenn  $\check{\psi}$  die skalare Helmholtzgleichung erfüllt. Mit:

$$\check{\vec{\mathcal{F}}}_2 = \frac{\vec{\nabla} \times \check{\vec{\mathcal{F}}}_1}{\check{k}} \quad (2.1.35)$$

erfüllt auch die zweite Hilfsfunktion  $\check{\vec{\mathcal{F}}}_2$  die HELMHOLTZ-Gleichung:

$$\vec{\nabla}^2 \check{\vec{\mathcal{F}}}_2 + \check{k}^2 \check{\vec{\mathcal{F}}}_2 = 0 \quad (2.1.36)$$

und es gilt der Zusammenhang:

$$\vec{\nabla} \times \check{\vec{\mathcal{F}}}_2 = \check{k} \check{\vec{\mathcal{F}}}_1 \quad (2.1.37)$$

Darüber hinaus ist auch  $\check{\vec{\mathcal{F}}}_2$  divergenzfrei. Bei der Konstruktion der Hilfsfunktionen war der Vektor  $\vec{c}$  zunächst beliebig, seine Wahl wird allerdings über die Symmetrieeigenschaft der Lösung entscheiden. Die Funktionen  $\check{\mathcal{F}}_{1/2}$  sind die Fundamentallösungen der HELMHOLTZ-Gleichung.

### Kugelkoordinaten - Multipole 3D

Um die HELMHOLTZ-Gleichungen in Kugelkoordinaten zu lösen, wählt man  $\vec{c} = \vec{e}_r$ . Dazu wird die skalare HELMHOLTZ-Gleichung in Kugelkoordinaten  $x, y, z \rightarrow r, \theta, \phi$  umgeschrieben:

$$\frac{1}{r^2} \partial_r (r^2 \partial_r \check{\psi}) + \frac{1}{r^2 \sin(\theta)} \partial_\theta (\sin(\theta) \partial_\theta \check{\psi}) + \frac{1}{r^2 \sin^2(\theta)} (\partial_\phi^2 \check{\psi}) + \check{k}^2 \check{\psi} = 0 \quad (2.1.38)$$

Mit dem Separationsansatz  $\check{\psi}(r, \theta, \phi) = \check{R}(r) \Theta(\theta) \Phi(\phi)$  erhält man die Lösung zu:

$$\check{\psi} = \begin{Bmatrix} \cos(m\phi) \\ \sin(m\phi) \end{Bmatrix} P_n^m(\cos(\theta)) \check{z}_n(\check{k}r) \quad (2.1.39)$$

Dabei sind die  $P_n^m$  die zugeordneten LEGENDRE-Funktionen der ersten Art mit dem Grad  $n$  und der Ordnung<sup>18</sup>  $m$  und die  $\check{z}_n$  sind die sphärischen BESSEL-Funktionen<sup>19</sup>. Der  $\cos(m\phi)$  beziehungsweise  $\sin(m\phi)$  Term beschreibt die gerade beziehungsweise ungerade Symmetrie bezüglich der  $\phi$ -Koordinate. Damit ergeben sich die Fundamentallösungen zu<sup>20</sup>:

$$\begin{aligned} \check{\vec{F}}_1 = & \frac{m}{\sin(\theta)} \begin{Bmatrix} -\sin(m\phi) \\ \cos(m\phi) \end{Bmatrix} P_n^m(\cos(\theta)) \check{z}_n(\check{k}r) \vec{e}_\theta \\ & - \begin{Bmatrix} \cos(m\phi) \\ \sin(m\phi) \end{Bmatrix} \frac{dP_n^m(\cos(\theta))}{d\theta} \check{z}_n(\check{k}r) \vec{e}_\phi \end{aligned} \quad (2.1.40)$$

und

$$\begin{aligned} \check{\vec{F}}_2 = & \frac{\check{z}_n(\check{k}r)}{\check{k}r} \begin{Bmatrix} \cos(m\phi) \\ \sin(m\phi) \end{Bmatrix} n(n+1) P_n^m(\cos(\theta)) \vec{e}_r \\ & + \begin{Bmatrix} \cos(m\phi) \\ \sin(m\phi) \end{Bmatrix} \frac{dP_n^m(\cos(\theta))}{d\theta} \frac{1}{(\check{k}r)} \frac{d}{d(\check{k}r)} [(\check{k}r) \check{z}_n(\check{k}r)] \vec{e}_\theta \\ & + m \begin{Bmatrix} -\sin(m\phi) \\ \cos(m\phi) \end{Bmatrix} \frac{P_n^m(\cos(\theta))}{\sin(\theta)} \frac{1}{(\check{k}r)} \frac{d}{d(\check{k}r)} [(\check{k}r) \check{z}_n(\check{k}r)] \vec{e}_\phi \end{aligned} \quad (2.1.41)$$

Die  $E$ - und  $H$  Felder erhält man schließlich mit:

$$\check{\vec{E}} = E_0 \check{\vec{F}}_1 \quad \check{\vec{H}} = -iY E_0 \check{\vec{F}}_2 \quad (2.1.42)$$

$$\check{\vec{H}} = H_0 \check{\vec{F}}_1 \quad \check{\vec{E}} = iZ H_0 \check{\vec{F}}_2 \quad (2.1.43)$$

Je nachdem, ob die Fundamentallösung  $\check{\vec{F}}_1$  das  $E$ - oder das  $H$ -Feld erzeugt, kann man hier von  $E$ - beziehungsweise  $H$ -polarisierten Wellen sprechen. Des Weiteren nimmt man noch eine Einteilung in gerade (jeweils obere Auswahl) und ungerade (jeweils untere Auswahl) vor. Eine Besonderheit ist, dass die mit  $\check{\vec{F}}_1$  assoziierte Feldkomponente keinen Anteil in radialer Richtung besitzt, das heißt  $\check{\vec{F}}_1 \cdot \vec{r} = 0$ .

Für den Fall  $\check{z}_n = \check{j}_n$  beschreiben die Lösungen stehende Kugelwellen. Sie bilden da-

<sup>18</sup>Mit  $m = 0$  entsprechen die  $P_n^m$  den LEGENDRE-Polynomen.

<sup>19</sup>Es gibt insgesamt vier sphärische BESSEL-Funktionen:

- $\check{j}_n$  die sphärische BESSEL-Funktion (im engeren Sinn)
- $\check{y}_n$  die sphärische NEUMANN-Funktion
- $\check{h}_n^{1,2} = \check{j}_n \pm i\check{y}_n$  die sphärischen HANKEL-Funktionen

<sup>20</sup>Häufig wird  $\check{\rho} = \check{k}r$  gesetzt.

mit die Eigenfunktionen von sphärischen Kavitäten (siehe Abschnitt 4.3). Wird hingegen  $\check{z}_n = \check{h}_n^{1,2}$  gewählt, beschreiben die resultierenden Felder ein- beziehungsweise auslaufende Kugelwellen. Diese entsprechen dann den bekannten sphärischen Multipolen. Für weitere Informationen siehe [14].

### Zylinderkoordinaten - Multipole 2D

Um die Lösung der MAXWELL-Gleichungen in Zylinderkoordinaten zu erhalten, wählt man den Vektor  $\vec{c} = \vec{e}_z$ . Analog zum Vorgehen in Kugelkoordinaten wird nun die skalare HELMHOLTZ-Gleichung in Zylinderkoordinaten  $x, y, z \rightarrow r, \varphi, z$  umgeschrieben:

$$\frac{1}{r} \partial_r (r \partial_r \check{\psi}) + \frac{1}{r^2} \partial_\phi^2 \check{\psi} + \partial_z^2 \check{\psi} + k^2 \check{\psi} = 0 \quad (2.1.44)$$

Mit Separationsansatz folgt<sup>21</sup>:

$$\check{\psi}(r, \phi, z) = \check{Z}_n(\check{k}r) e^{in\phi} \quad (2.1.45)$$

Dabei ist  $\check{Z}_n$  die Besselfunktion<sup>22</sup>. Damit ergeben sich die (vereinfachten) Fundamentallösungen zu:

$$\check{\vec{F}}_1 = \check{k} \left( in \frac{\check{Z}_n(\check{k}r)}{(\check{k}r)} \vec{e}_r - \frac{d}{d(\check{k}r)} \check{Z}_n(\check{k}r) \vec{e}_\phi \right) e^{in\phi} \quad (2.1.46)$$

und

$$\check{\vec{F}}_2 = \check{k} \check{Z}_n(\check{k}r) \vec{e}_z e^{in\phi} \quad (2.1.47)$$

Die Konstruktion der  $E$ - und  $H$ -Felder erfolgt wiederum mit den Gleichungen 2.1.42 und 2.1.43. Insbesondere kann man auch hier wieder, je nachdem ob die Fundamentallösung  $\check{\vec{F}}_1$  das  $E$ - oder das  $H$ -Feld erzeugt, von  $E$ - beziehungsweise  $H$ -polarisierten Wellen sprechen. Dabei tritt der Fall auf, dass die  $H$ -Polarisation identisch mit dem TM-Fall der zweidimensionalen MAXWELL-Gleichungen ist, das heißt aus den Feldkomponenten  $H_x$ ,  $H_y$  und  $E_z$  besteht, während die  $E$ -Polarisation identisch mit dem TE-Fall ist, welcher aus den Feldkomponenten  $E_x$ ,  $E_y$  und  $H_z$  besteht.

<sup>21</sup>Hier: Vereinfachte Lösung ohne Propagation in  $z$ -Richtung. Für die allgemeine Lösung siehe [4], Kapitel 8.4.

<sup>22</sup>Es gibt insgesamt vier BESSEL-Funktionen:

- $\check{J}_n$  die BESSEL-Funktion (1. Art)
- $\check{Y}_n$  die NEUMANN-Funktion (auch BESSEL-Funktion 2. Art)
- $\check{H}_n^{1,2} = \check{J}_n \pm i\check{Y}_n$  die HANKEL-Funktionen

Für den Fall  $\check{Z}_n = \check{J}_n$  beschreiben die Lösungen stehende Zylinderwellen. Sie bilden damit die Eigenfunktionen von zylindrischen Kavitäten (siehe Abschnitt 4.2). Wird hingegen  $\check{Z}_n = \check{H}_n^{1,2}$  gewählt, beschreiben die resultierenden Felder ein- beziehungsweise auslaufende Zylinderwellen. Diese entsprechen dann den bekannten zylindrischen Multipolen. Für weitere Informationen siehe [14].

## 2.2. EULER - Gleichung

Als zweites wichtiges PDGL-System neben den MAXWELL-Gleichungen werden hier die EULER-Gleichungen behandelt. Diese beschreiben die Bewegung eines reibungsfreien (nicht-viskosen) Fluids ([32]), wie man es unter anderem auch für Elektronengase annimmt. Die EULER-Gleichungen in Erhaltungssatzform lauten:

$$\partial_t \rho + \vec{\nabla} \rho \vec{v} = 0 \quad (2.2.1a)$$

$$\partial_t \rho \vec{v} + \vec{\nabla}(\rho \vec{v} \otimes \vec{v} + p) = 0 \quad (2.2.1b)$$

$$\partial_t U + \vec{\nabla}[\vec{v}(U + p)] = 0 \quad (2.2.1c)$$

Dabei ist  $\rho$  die Dichte,  $\vec{v}$  die Geschwindigkeit,  $p$  der Druck und  $U$  die gesamte Energie pro Volumen des Fluids. Dabei beschreibt die erste Gleichung die Erhaltung der Dichte (und damit Masse, Ladung, Anzahl, ... der zugrunde liegenden Teilchenart), die zweite die Erhaltung des Impulses und die dritte Gleichung die Erhaltung der Energie. Auch dieser Gleichungssatz ist analog zu den MAXWELL-Gleichungen noch nicht vollständig. Es fehlt noch die Verknüpfung von Druck, Dichte, Geschwindigkeit und Energie, welche durch die zugehörige Zustandsgleichung beschrieben wird.

### 2.2.1. Innere Energie und Druck für nichtviskose Gase

Als Referenz für die Zustandsgleichung sei hier diejenige eines klassischen Fluids angegeben:

$$U = \underbrace{\frac{p}{\gamma - 1}}_{\text{innere Energie}} + \underbrace{\frac{1}{2} \rho \vec{v}^2}_{\text{kinetische Energie}} \quad (2.2.2)$$

Dabei ist  $\gamma = c_p/c_v$  der adiabatische Exponent,  $c_p$  die spezifische Wärmekapazität bei konstantem Druck und  $c_v$  die spezifische Wärmekapazität bei konstantem Volumen. Bei einatomigen Gasen ist  $\gamma = 1.4$ . Stellt man diese Gleichung nach dem Druck um, erhält man:

$$p = (\gamma - 1) \left( U - \frac{1}{2} \rho \vec{v}^2 \right) \quad (2.2.3)$$

### 2.2.2. Innere Energie und Druck für Elektronengase

Im Vergleich zu klassischen Fluiden gestaltet sich die Berechnung der Zustandsgleichung an dieser Stelle etwas aufwändiger, da hier noch das PAULI-Prinzip berücksichtigt werden muss. Dazu benötigt man das Volumen der  $k$ -Raumkugel  $V_k$  im Impulsraum der Elektronen:

$$V_k = g_d k^{\alpha_d} \quad (2.2.4)$$

Dabei ist  $\alpha_d$  die Dimension und  $g_d$  ein dimensionsabhängiger Faktor (für den 3D:  $g_d = \frac{4}{3}\pi$ ). Des Weiteren ist das Volumen eines Zustandes im  $k$ -Raum in Abhängigkeit von dem Volumen  $V$  des Körpers:

$$\Omega_k = \frac{(2\pi)^{\alpha_d}}{V} \quad (2.2.5)$$

Damit ergibt sich die Anzahl der Elektronen zu:

$$N_k = \frac{V_k}{\Omega_k} = g_s \frac{g_d}{(2\pi)^{\alpha_d}} V k^{\alpha_d} \quad (2.2.6)$$

Dabei wurde berücksichtigt, dass jeder Zustand mehrfach besetzt werden kann (bei Elektronen auf Grund des Spins zweifach:  $g_s = 2$ ). Mit Hilfe der Dispersionsrelation:

$$u = g_p k^{\alpha_p} \quad (2.2.7)$$

lässt sich noch ein Ausdruck für die Energie  $u$  gewinnen. Dabei ist für freie Elektronen  $\alpha_p = 2$  und  $g_p = \frac{\hbar^2}{2m}$ . Damit berechnet sich die Anzahl der Elektronen als Funktion der Energie zu:

$$N_k = \frac{g_s}{(2\pi)^{\alpha_d}} \frac{g_d}{g_p^{\frac{\alpha_d}{\alpha_p}}} V u^{\frac{\alpha_d}{\alpha_p}} \quad (2.2.8)$$

Nun lässt sich die Zustandsdichte  $D_k$  berechnen:

$$D_k = \partial_u N_k = \frac{g_s}{(2\pi)^{\alpha_d}} \frac{g_d}{g_p^{\frac{\alpha_d}{\alpha_p}}} V \frac{\alpha_d}{\alpha_p} u^{\frac{\alpha_d - \alpha_p}{\alpha_p}} \quad (2.2.9)$$

Und mit ihr die Energie  $U$ :

$$U = \int_0^{u_F} u D_k(u) du = \frac{g_s}{(2\pi)^{\alpha_d}} \frac{g_d}{g_p^{\frac{\alpha_d}{\alpha_p}}} V \frac{\alpha_d}{\alpha_p} \frac{\alpha_p}{\alpha_d + \alpha_p} u_F^{\frac{\alpha_d + \alpha_p}{\alpha_p}} \quad (2.2.10)$$



Dabei ist die FERMI-Energie mit Gleichung 2.2.8:

$$u_F = g_p \left( \frac{(2\pi)^{\alpha_d} N}{g_s g_d V} \right)^{\frac{\alpha_p}{\alpha_d}} \quad (2.2.11)$$

Einsetzen in Gleichung 2.2.10 liefert:

$$\begin{aligned} U &= \frac{g_s g_d}{(2\pi)^{\alpha_d} g_p^{\frac{\alpha_d}{\alpha_p}}} V \frac{\alpha_d}{\alpha_d + \alpha_p} g_p^{\frac{\alpha_d + \alpha_p}{\alpha_p}} \left( \frac{(2\pi)^{\alpha_d} N}{g_s g_d V} \right)^{\frac{\alpha_d + \alpha_p}{\alpha_d}} \\ &= g_p \left( \frac{(2\pi)^{\alpha_d}}{g_s g_d} \right)^{\frac{\alpha_p}{\alpha_d}} \frac{\alpha_d}{\alpha_d + \alpha_p} N \left( \frac{N}{V} \right)^{\frac{\alpha_p}{\alpha_d}} \end{aligned} \quad (2.2.12)$$

Damit lässt sich jetzt der Druck  $p$  berechnen zu:

$$p = -(\partial_V U)_N = g_p \left( \frac{(2\pi)^{\alpha_d}}{g_s g_d} \right)^{\frac{\alpha_p}{\alpha_d}} \frac{\alpha_p}{\alpha_d + \alpha_p} \left( \frac{N}{V} \right)^{\frac{\alpha_p + \alpha_d}{\alpha_d}} \quad (2.2.13)$$

Ersetzt man jetzt noch  $N/V$  mit der Teilchendichte  $n$  und bezeichnet den Vorfaktor mit  $\zeta$ , ergibt sich der THOMAS-FERMI-Druck für freie Elektronengase:

$$p = \zeta(\alpha_d, \alpha_p) n^{\frac{\alpha_p + \alpha_d}{\alpha_d}} \quad (2.2.14)$$

Dabei wird deutlich, dass der Druck nicht mehr, wie bei klassischen Fluiden, von der Energie abhängt. Damit entkoppeln sich die Gleichungen für Energie- und Impulserhaltung, so dass erstere zur Lösung nicht mehr benötigt wird. Für den dreidimensionalen Raum erhält man für das freie Elektronengas schließlich:

$$p = \frac{\hbar^2}{5m} (3\pi^2)^{\frac{2}{3}} n^{\frac{5}{3}} \quad (2.2.15)$$

Dies entspricht auch genau dem Ausdruck der in [39, 61] verwendeten Formel. Diese EULER-Gleichungen werden in Abschnitt 2.3.3 zu einem Materialmodell zur Beschreibung der nichtlokalen, nichtlinearen Eigenschaften von Metallen erweitert.

## 2.3. Materialmodelle

Nachdem in den ersten beiden Abschnitten die beiden wichtigen partiellen Differentialgleichungssysteme (PDGL) behandelt wurden, geht es nun um die Beschreibung der für die Simulation wichtigen Materialmodelle ( $\vec{P}(\vec{r}, t)$ ). Diese bestehen aus dem DRUDE-LORENTZ-Modell, einer Antireflexschicht (engl.: Perfectly Matched Layer: PML) und dem Hydrodynamischen Elektronengasmodell.

### 2.3.1. DRUDE-LORENTZ Modell

Das für die Simulation von nanooptischen Systemen wichtigste Materialmodell (neben dem trivialen Fall des konstanten Brechungsindex) ist das DRUDE-LORENTZ-Modell. Mit diesem lassen sich nahezu alle linearen lokalen Materialien beschreiben. Eingehendere Betrachtungen dazu finden sich unter anderem in [4]. Hier erfolgt nur die Herleitung des LORENTZ-Modells. Das DRUDE-Modell folgt dann daraus.

Zur Herleitung betrachtet man die gedämpfte Bewegung einer Ladung in einem harmonischen Potential bei angelegtem externen elektromagnetischen Feld:

$$m\ddot{\vec{r}} = -d\dot{\vec{r}} - k\vec{r} + \underbrace{q \left( \vec{E} + \mu\dot{\vec{r}} \times \vec{H} \right)}_{\text{LORENTZ-Kraft}} \quad (2.3.1)$$

Dabei beschreibt  $m$  die Masse der Ladung,  $d$  die Dämpfung,  $k$  die Federkonstante,  $q$  die Ladung,  $\vec{r}$  den Ort der Ladung und  $\vec{E}$  bzw.  $\vec{H}$  das elektromagnetische Feld. Als nächstes wird angenommen, dass sowohl  $\vec{E}$  wie auch  $\vec{H}$ -Feld im Bereich der Amplitude der Ladung nur schwach variieren. Damit kann man  $\vec{E}$  und  $\vec{H}$  mittels Taylorentwicklung um den Ruhepunkt der Ladung nähern<sup>23</sup>:

$$\vec{E} \approx \vec{E}_0 + \left( \vec{\nabla} \otimes \vec{E} \right)_0 \vec{r} \quad (2.3.2a)$$

$$\dot{\vec{r}} \times \vec{H} \approx \dot{\vec{r}} \times \vec{H}_0 \quad (2.3.2b)$$

Setzt man nun die Eigenfrequenz  $\omega_0 = \sqrt{\frac{k}{m}}$ , die Dämpfungskonstante  $\gamma = \frac{d}{m}$ , die Polarisation  $\vec{P} = N\vec{p}$ , das Dipolmoment  $\vec{p} = q\vec{r}$ , die Elektronendichte  $N$  und die Plasmafrequenz

<sup>23</sup>Dabei werden hier die Komponenten der LORENTZ-Kraft, bei Annahme einer harmonischen Bewegung, bis zur zweiten Ordnung in  $\omega$  berücksichtigt.

$\omega_p = \sqrt{\frac{Nq^2}{m\epsilon_0}}$  ein, erhält man:

$$\partial_t^2 \vec{P} = \omega_p^2 \epsilon_0 (\vec{E} + \frac{\mu}{Nq} \partial_t \vec{P} \times \vec{H}) - \gamma \partial_t \vec{P} - \omega_0^2 \vec{P} \quad (2.3.3a)$$

$$\partial_t^2 \vec{P} \approx \underbrace{\omega_p^2 \epsilon_0 \vec{E}_0 - \gamma \partial_t \vec{P} - \omega_0^2 \vec{P}}_{\text{lineares LORENTZ Modell}} + \underbrace{\frac{q}{m} (\vec{\nabla} \otimes \vec{E})_0 \vec{P} + \frac{\mu q}{m} \partial_t \vec{P} \times \vec{H}}_{\text{nichtlineare Erweiterung 2. Ordnung}} \quad (2.3.3b)$$

Den Übergang zur Frequenzdomäne erhält man mit der Annahme  $\vec{P}(t) = \Re(\check{\vec{P}}_0 e^{-i\omega t})$ . Einsetzen in Gleichung 2.3.3b bei Vernachlässigen des nichtlinearen Anteils und Umstellen nach  $\check{\vec{P}}$  liefert unter Vernachlässigung des nichtlinearen Anteils:

$$\check{\vec{P}}(\omega) = \frac{\omega_p^2}{\underbrace{\omega_0^2 - \omega^2 - i\gamma\omega}_{\check{\chi}_E(\omega)}} \epsilon_0 \check{\vec{E}} \quad (2.3.4)$$

das bekannte LORENTZ-Modell. Das DRUDE-Modell folgt im Grenzwert einer vernachlässigbaren Rückstellkraft, d.h für  $\omega_0 = 0$ . Nach den Bemerkungen in den Abschnitten 2.1 und 2.1.2 (Gleichung 2.1.3 und 2.1.11) lassen sich für komplexere Materialeigenschaften, bei denen ein einzelner LORENTZ-Pol nicht zur Beschreibung genügt, mehrere dieser addieren, d.h.:

$$\begin{aligned} \vec{P} &= \epsilon_0 \chi_E^\infty \vec{E} + \sum_i \vec{P}_i & \partial_t^2 \vec{P}_i &= \omega_{p,i}^2 \epsilon_0 \vec{E}_0 - \gamma_i \partial_t \vec{P}_i - \omega_{0,i}^2 \vec{P}_i \\ \check{\chi}_E &= \chi_E^\infty + \sum_i \check{\chi}_{E,i} & \check{\chi}_{E,i} &= \frac{\omega_{p,i}^2}{\omega_{0,i}^2 - \omega^2 - i\gamma_i \omega} \end{aligned}$$

Für die Simulation der in dieser Arbeit verwendeten Metalle wurde ein Modell aus konstantem, einem DRUDE- und einem LORENTZ-Term verwendet:

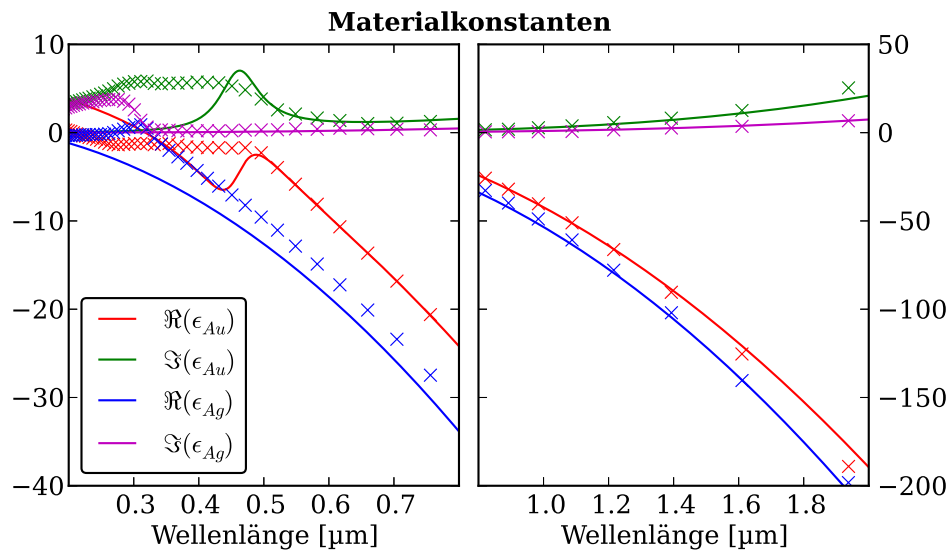
$$\check{\epsilon}_r(\omega) = \epsilon^\infty + \underbrace{\frac{-\omega_{p,D}^2}{\omega^2 + i\gamma_D \omega}}_{\text{DRUDE-Pol}} + \underbrace{\frac{\omega_{p,L}^2}{\omega_{0,L}^2 - \omega^2 - i\gamma_L \omega}}_{\text{LORENTZ-Pol}} \quad (2.3.5)$$

Die Werte der Materialparameter sind für Silber (aus [46]) und Gold (aus [65]) in Tabelle 2.3.1 zusammengefasst. Der Verlauf der dielektrischen Funktion ist in Abbildung 2.3.1 dargestellt. Dass sich das LORENTZ-Modell auch für Approximationen komplizierterer Materialien eignet, zeigt die Abbildung 2.3.2. In dieser sind als Beispiel die gemessenen und die mit drei LORENTZ-Polen angenäherten Materialkonstanten für Zinkphthalocyanin (ZnPC, ein in der organischen Elektronik häufig verwendetes Molekül) dargestellt. Die numerischen

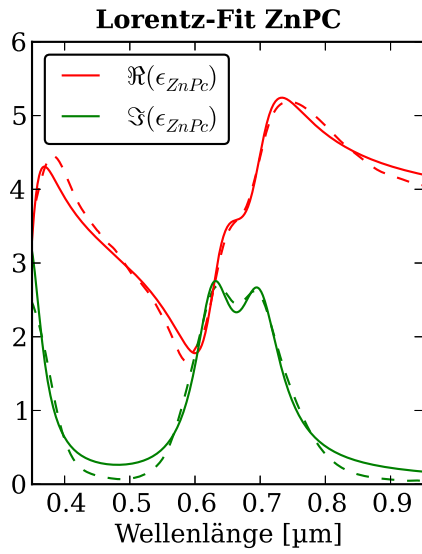
**Tabelle 2.3.1.:** Verwendete Materialparameter von Silber und Gold nach [46] und [65].

	Silber	Gold
$\epsilon^\infty$	1.0	5.9673
$\omega_{p,D}$	$1.39 \cdot 10^{16} \text{ 1/s}$	$1.3280 \cdot 10^{16} \text{ 1/s}$
$\gamma_D$	$3.23 \cdot 10^{13} \text{ 1/s}$	$1.0003 \cdot 10^{14} \text{ 1/s}$
$\omega_{0,L}$	-	$4.0845 \cdot 10^{15} \text{ 1/s}$
$\omega_{p,L}$	-	$4.2644 \cdot 10^{15} \text{ 1/s}$
$\gamma_L$	-	$6.5885 \cdot 10^{14} \text{ 1/s}$

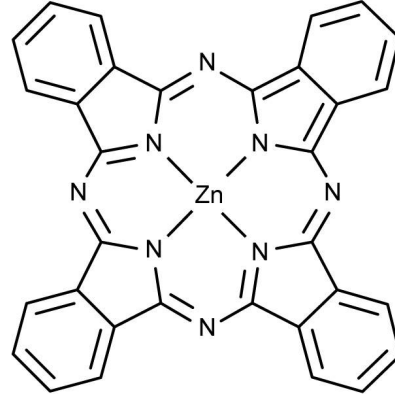
Werte der ermittelten LORENTZ-Pole sind in Tabelle 2.3.2 aufgeführt.



**Abbildung 2.3.1.:** Dielektrische Funktion der verwendeten Metalle ( $\epsilon_r$ ) im Wellenlängenbereich von 0.2 – 0.8 μm (links) und 0.8 – 2.0 μm (rechts). Die Kreuze sind die experimentellen Daten nach [22], die Linien die angenäherten Daten nach [46] und [65].



(a) Dielektrische Funktion ( $\epsilon_r$ ) von ZnPC.



(b) Strukturformel von ZnPC

**Abbildung 2.3.2.:** LORENTZ-Fit von ZnPC. Die gestrichelten Linien stellen die gemessenen, die durchgezogenen die mit drei LORENTZ-Polen angenäherten Werte dar.

**Tabelle 2.3.2.:** LORENTZ-Fit von ZnPC (Vergleiche Abbildung 2.3.2).

	$\epsilon^\infty$	$\omega_{0,L}$	$\omega_{p,L}$	$\gamma_L$
-	2.785	-	-	-
1. Pol	-	$5.463 \cdot 10^{15} \text{ 1/s}$	$3.555 \cdot 10^{15} \text{ 1/s}$	$7.090 \cdot 10^{14} \text{ 1/s}$
2. Pol	-	$3.002 \cdot 10^{15} \text{ 1/s}$	$1.512 \cdot 10^{15} \text{ 1/s}$	$3.357 \cdot 10^{14} \text{ 1/s}$
3. Pol	-	$2.699 \cdot 10^{15} \text{ 1/s}$	$1.295 \cdot 10^{15} \text{ 1/s}$	$3.010 \cdot 10^{14} \text{ 1/s}$

### 2.3.2. PML - Antireflexschicht

Ein zweites wichtiges Materialmodell, für die Numerik ist die der Antireflexschicht (engl.: Perfectly Matched Layer: PML). Diese wird dazu verwendet, um aus einem Simulationsgebiet herauslaufende Wellen ohne Reflexion zu absorbieren, und so bei endlichem Simulationsgebiet eine reale Wellenausbreitung in den unbegrenzten Freiraum abzubilden (Näheres zur Anwendung siehe Abschnitt 2.6.2). Erstmals beschrieben wurde sie in [2]. Adaptionen für die DG-Methode finden sich unter anderem in [7, 20, 35, 45, 46]. Hier wird dabei in Anlehnung an die angegebenen Quellen nur der uniaxiale PML behandelt. Eine andere Version, den „Stretched Coordinate PML“, der vor allem bei dispersiven Materialien im PML-Gebiet zum Einsatz kommt, findet man in [27, 29].

Da das Material, wie bereits erwähnt, das elektromagnetische Feld absorbieren soll, beginnt man mit einer Erweiterung der MAXWELL-Gleichungen um einen Dämpfungsterm:

$$\epsilon \partial_t \vec{E} = -\epsilon \overset{\leftrightarrow}{\gamma}_E \vec{E} + \vec{\nabla} \times \vec{H} \quad (2.3.6a)$$

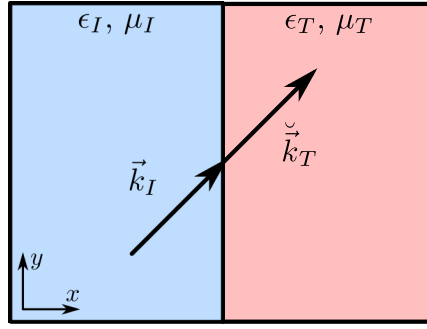
$$\mu \partial_t \vec{H} = -\mu \overset{\leftrightarrow}{\gamma}_H \vec{H} - \vec{\nabla} \times \vec{E} \quad (2.3.6b)$$

Dabei ist  $\overset{\leftrightarrow}{\gamma}$  ein Diagonal-Tensor mit der Einheit  $1/s$ . Dies entspricht einer richtungsabhängigen Dämpfungskonstante. Dabei ist  $\epsilon \overset{\leftrightarrow}{\gamma}$  gleich der spezifischen Leitfähigkeit  $\overset{\leftrightarrow}{\sigma}$ . Mit dem gewählten Ansatz ist jedoch der Übergang zur Frequenzdomäne etwas einfacher:

$$\underbrace{-i\omega\epsilon \left(1 + i \frac{\overset{\leftrightarrow}{\gamma}_E}{\omega}\right)}_{\overset{\leftrightarrow}{s}_E} \vec{E} = +\vec{\nabla} \times \vec{H} \quad (2.3.7a)$$

$$\underbrace{-i\omega\mu \left(1 + i \frac{\overset{\leftrightarrow}{\gamma}_H}{\omega}\right)}_{\overset{\leftrightarrow}{s}_H} \vec{H} = -\vec{\nabla} \times \vec{E} \quad (2.3.7b)$$

Für die weitere Herleitung wird nun der Übersichtlichkeit wegen nur die  $TE$ -Polarisation mit einem in  $x$ -Richtung absorbierendem PML behandelt (siehe Abbildung 2.3.3). Mit dieser Vereinfachung und der Annahme „ebener“ Wellen als Lösung lauten die MAXWELL-



**Abbildung 2.3.3.:** Absorption am PML, mit der einfallenden Welle  $I$  und der transmittierten (bzw. absorbierten) Welle  $T$ .

Gleichungen nun:

$$\omega \epsilon \check{s}_{E_x} \check{E}_x = -\check{k}_y \check{H}_z \quad (2.3.8a)$$

$$\omega \epsilon \check{s}_{E_y} \check{E}_y = \check{k}_x \check{H}_z \quad (2.3.8b)$$

$$\omega \mu \check{s}_{H_z} \check{H}_z = \check{k}_x \check{E}_y - \check{k}_y \check{E}_x \quad (2.3.8c)$$

Auflösen der ersten beiden Gleichungen 2.3.8 nach  $\check{E}_x$  bzw.  $\check{E}_y$ , sowie Einsetzen in die Dritte liefert die Dispersionsrelation:

$$\omega^2 \mu \epsilon \check{s}_{H_z} \check{s}_{E_y} \check{s}_{E_x} = \check{s}_{E_x} \check{k}_x^2 + \check{s}_{E_y} \check{k}_y^2 \quad (2.3.9)$$

Mit Hilfe der Stetigkeitsbedingung (Abschnitt 2.1.3) unter der Bedingung verschwindender Reflexion, das heißt,  $\epsilon_I \check{E}_{I,x} = \epsilon_T \check{E}_{T,x}$ ,  $\check{E}_{I,y} = \check{E}_{T,y}$  und  $\check{H}_{I,z} = \check{H}_{T,z}$ , folgt aus den Gleichungen 2.3.8a und 2.3.8b:

$$\frac{-\check{k}_{I,y} \check{H}_{I,z}}{\omega} = \frac{-\check{k}_{T,y} \check{H}_{T,z}}{\omega} \quad (2.3.10a)$$

$$\frac{\check{k}_{I,x} \check{H}_{I,z}}{\omega \epsilon_I} = \frac{\check{k}_{T,x} \check{H}_{T,z}}{\omega \epsilon_T \check{s}_{E_y}} \quad (2.3.10b)$$

Damit erhält man den Zusammenhang der  $k$ -Komponenten von einfallender und transmittierter Welle durch Vergleich der  $H_z$ -Komponenten:

$$\check{k}_{T,y} = \check{k}_{I,y} \quad (2.3.11a)$$

$$\check{k}_{T,x} = \frac{\epsilon_T}{\epsilon_I} \check{s}_{E_y} \check{k}_{I,x} \quad (2.3.11b)$$

Wählt man nun  $\epsilon_I = \epsilon_T$  und  $\mu_I = \mu_T$  und verwendet Gleichung 2.3.11 zusammen mit der

Dispersionrelation (Gleichung 2.3.9), ergibt sich eine Bedingung für die  $\check{s}_{E_x}$ ,  $\check{s}_{E_y}$  und  $\check{s}_{H_z}$ :

$$\check{s}_{E_x}\check{s}_{E_y}k_{I,x}^2 + k_{I,y}^2 = \check{s}_{H_z}\check{s}_{E_x}k_I^2 \quad (2.3.12a)$$

$$= \check{s}_{H_z}\check{s}_{E_x}(k_{I,x}^2 + k_{I,y}^2) \quad (2.3.12b)$$

Diese Gleichung lässt sich unter anderem mit folgendem Ansatz erfüllen:

$$\check{s}_x := \check{s}_{E_y} = \check{s}_{H_z} = \check{s}_{E_x}^{-1} \quad (2.3.13)$$

Dabei ist die konkrete Wahl von  $\check{s}_x$  beliebig, so dass man auch frei in der Wahl von  $\gamma_x$  ist. Damit schreiben sich der Permittivitätstensor und der sinngemäß konstruierte Permeabilitätstensor nun:

$$\overset{\leftrightarrow}{\epsilon} = \epsilon \begin{bmatrix} \check{s}_x^{-1} & 0 & 0 \\ 0 & \check{s}_x & 0 \\ 0 & 0 & \check{s}_x \end{bmatrix} \quad \text{bzw.} \quad \overset{\leftrightarrow}{\mu} = \mu \begin{bmatrix} \check{s}_x^{-1} & 0 & 0 \\ 0 & \check{s}_x & 0 \\ 0 & 0 & \check{s}_x \end{bmatrix} \quad (2.3.14)$$

Der noch notwendige Übergang von der Absorption in einer Dimension zu der Absorption in drei Dimensionen erfolgt analog zu [7, 45, 46]. Dabei multipliziert man einfach die erhaltenen Tensoren (hier am Beispiel von  $\overset{\leftrightarrow}{\epsilon}$ ):

$$\overset{\leftrightarrow}{\epsilon} = \epsilon \begin{bmatrix} \check{s}_x^{-1} & 0 & 0 \\ 0 & \check{s}_x & 0 \\ 0 & 0 & \check{s}_x \end{bmatrix} \cdot \begin{bmatrix} \check{s}_y & 0 & 0 \\ 0 & \check{s}_y^{-1} & 0 \\ 0 & 0 & \check{s}_y \end{bmatrix} \cdot \begin{bmatrix} \check{s}_z & 0 & 0 \\ 0 & \check{s}_z & 0 \\ 0 & 0 & \check{s}_z^{-1} \end{bmatrix} = \epsilon \begin{bmatrix} \frac{\check{s}_y\check{s}_z}{\check{s}_x} & 0 & 0 \\ 0 & \frac{\check{s}_x\check{s}_z}{\check{s}_y} & 0 \\ 0 & 0 & \frac{\check{s}_x\check{s}_y}{\check{s}_z} \end{bmatrix} \quad (2.3.15)$$

Nun steht noch die Rücktransformation in die Zeitdomäne an. Dazu wird im Folgenden wieder der Übersichtlichkeit wegen nur die  $\check{E}_x$ -Komponente behandelt. Die Gleichung liest sich mit  $\vec{\nabla} \times \vec{H} = \dots$ :

$$\check{\omega}\epsilon\check{E}_x = -\check{\omega}\epsilon\left(\frac{\check{s}_y\check{s}_z}{\check{s}_x} - 1\right)\check{E}_x + \dots \quad (2.3.16)$$

Nun führt man Hilfsstromdichten  $\check{J}_x$  und Hilfspolarisationen  $\check{P}_x$  ein:

$$\check{J}_x = \check{\omega}\epsilon\left(\frac{\check{s}_y\check{s}_z}{\check{s}_x} - 1\right)\check{E}_x \quad (2.3.17a)$$

$$= \frac{\check{\omega}\epsilon}{\check{\omega} + \gamma_x}\left(\gamma_y + \gamma_z - \gamma_x + \frac{\gamma_y\gamma_z}{\check{\omega}}\right)\check{E}_x \quad (2.3.17b)$$

$$\check{P}_x = \check{J}_x - \epsilon(\gamma_z + \gamma_y - \gamma_x)\check{E}_x \quad (2.3.17c)$$



Einsetzen von Gleichung 2.3.17b in Gleichung 2.3.17c liefert:

$$\check{\omega}\check{P}_x = -\gamma_x\check{P}_x + \epsilon(\gamma_x\gamma_x + \gamma_y\gamma_z - \gamma_y\gamma_x - \gamma_z\gamma_x)\check{E}_x \quad (2.3.18)$$

Dieser Ausdruck lässt sich nun leicht in die Zeitdomäne transformieren:

$$\partial_t E_x = -P_x - \epsilon(\gamma_z + \gamma_y - \gamma_x)E_x + \dots \quad (2.3.19a)$$

$$\partial_t P_x = -\gamma_x P_x + \epsilon(\gamma_x\gamma_x + \gamma_y\gamma_z - \gamma_y\gamma_x - \gamma_z\gamma_x)E_x \quad (2.3.19b)$$

In Abschnitt 3.2ff. werden diese Gleichungen hinsichtlich ihrer Effizienz untersucht und für die Konstruktion des PMLs verwendet.

### 2.3.3. Hydrodynamisches Elektronengas

Neben den klassischen linearen lokalen Materialien, wie sie mittels des DRUDE-LORENTZ-Modells beschrieben werden können, soll in dieser Arbeit auch speziell auf die Erzeugung höherer Harmonischer mittels Metallnanostrukturen eingegangen werden. Da in Metallen die Hauptbeiträge von den quasifreien Leitungselektronen her rühren, muss darauf der Schwerpunkt der Modellierung gelegt werden, zum Beispiel mit dem Modell des hydrodynamischen Elektronengases. Dieses beschreibt die Elektronen als Fluid in einem Potentialtopf mit unendlich hohen Wänden am Rand der Metallstruktur. Anwendungen dieses Ansatzes finden sich z. B. in [9, 39, 40, 61, 73]. Als Bewegungsgleichung werden in diesem Abschnitt die EULER-Gleichungen für Elektronen verwendet. Dabei schreiben sich die Größen Ladungsdichte  $\rho$ , Stromdichte  $\vec{j}$  mittels der Teilchendichte für Elektronen und positiven Ionen  $n^-$  bzw.  $n^+$ , der Elementarladung  $e_0$  und der Geschwindigkeit der Elektronen  $\vec{v}$  (dabei sind die positiven Ionen als unbeweglich angenommen):

$$\rho = \rho^+ + \rho^- \quad (2.3.20a)$$

$$\vec{j} = -e_0 n^- \vec{v} \quad (2.3.20b)$$

$$\rho^+ = e_0 n^+ \quad (2.3.20c)$$

$$\rho^- = -e_0 n^- \quad (2.3.20d)$$

Die Kopplung der EULER- mit den MAXWELL-Gleichungen erfolgt über die Stromdichte bzw. die LORENTZ-Kraft. Damit liest sich der komplette Gleichungssatz aus EULER- und MAXWELL-Gleichungen, unter Beachtung der Entkopplung von der Energie  $U$  und dem

Druck  $p = p(n^-)$  zu:

$$\partial_t n^- + \nabla \cdot (n^- \vec{v}) = 0 \quad (2.3.21a)$$

$$\partial_t (n^- \vec{v}) + \nabla \cdot \left( n^- \vec{v} \otimes \vec{v} + \frac{1}{m_e} p \right) = \frac{-e_0}{m_e} n^- \left( \vec{E} + \mu \vec{v} \times \vec{H} \right) - \frac{1}{\tau} n^- \vec{v} \quad (2.3.21b)$$

$$\epsilon \partial_t E = \vec{\nabla} \times \vec{H} - (-e_0 n^- \vec{v}) \quad (2.3.21c)$$

$$\mu \partial_t H = -\vec{\nabla} \times \vec{E} \quad (2.3.21d)$$

Dabei wurde noch die effektive Elektronenmasse  $m_e$  und der „Quanten“-Druck  $p$  in THOMAS-FERMI-Näherung (siehe Gleichung 2.2.15) benutzt.

$$p = \frac{\hbar^2}{5m_e} (3\pi^2)^{2/3} (n^-)^{5/3}$$

## 2.4. Plasmonik

Im Gegensatz zu nichtleitenden Dielektrika, die im wesentlichen über gebundene Elektronen verfügen, besitzen Metalle eine große Anzahl quasifreier Elektronen im Leitungsband. Dieser Unterschied führt dazu, dass die Antwort eines Dielektrika auf ein angelegtes Lichtfeld im wesentlichen aus einer Ansammlung induzierter Dipole besteht, welche schnell in die Ruhelage zurückpendeln. Dem gegenüber kann es bei Metallen gelingen, die freien Elektronen zu einer kollektiven Schwingung anzuregen. Diese wird Plasmon genannt. Ein umfassender Überblick über die Eigenschaften von Plasmonen findet sich in [58]. Im Folgenden wird kurz auf die wichtigsten Phänomene und Begriffe eingegangen.

### 2.4.1. Filme

Einer der einfachsten Zugänge zum Thema Plasmonen besteht in der Eigenlösung der Grenzfläche nach Abschnitt 2.1.4. Mit  $\check{\mu}_1 = \check{\mu}_2 = \check{\mu}$  lässt sich die Gleichung 2.1.30 noch weiter vereinfachen:

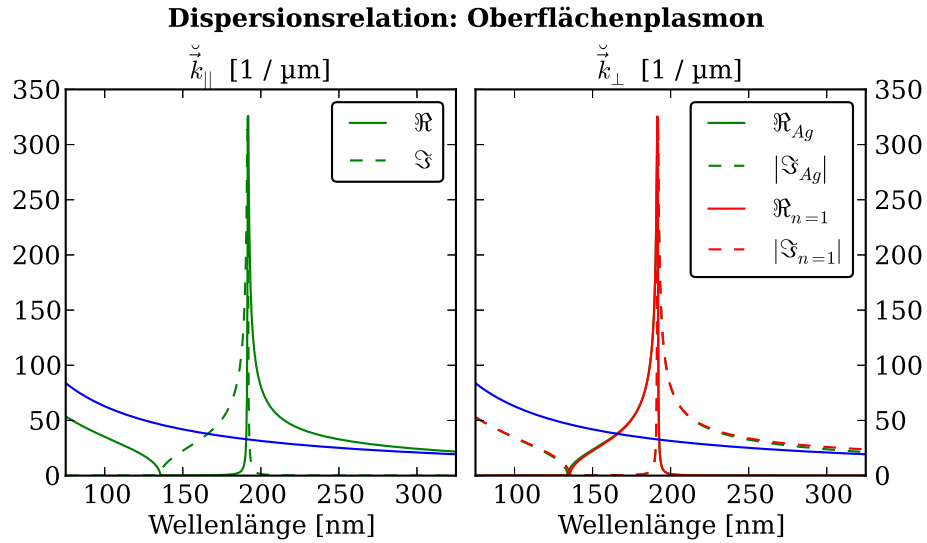
$$\check{k}_{\parallel}^2 = \omega^2 \check{\mu} \frac{\check{\epsilon}_1 \check{\epsilon}_2}{(\check{\epsilon}_1 + \check{\epsilon}_2)} \quad (2.4.1)$$

Da diese Plasmonen an einer Dielektrika-Metall-Oberfläche entlanglaufen, werden sie häufig als Oberflächen-Plasmon-Polariton (SPP von engl.: „Surface Plasmon Polariton“) bezeichnet. Mit Hilfe von Gleichung 2.1.19 lässt sich noch die Plasmonwellenlänge  $\lambda_{SPP}$  berechnen:

$$\check{\lambda}_{SPP} = \frac{2\pi}{\check{k}_{\parallel}} \quad (2.4.2)$$

Die notwendigen  $\check{k}_{\perp}$  Komponenten ergeben sich aus:

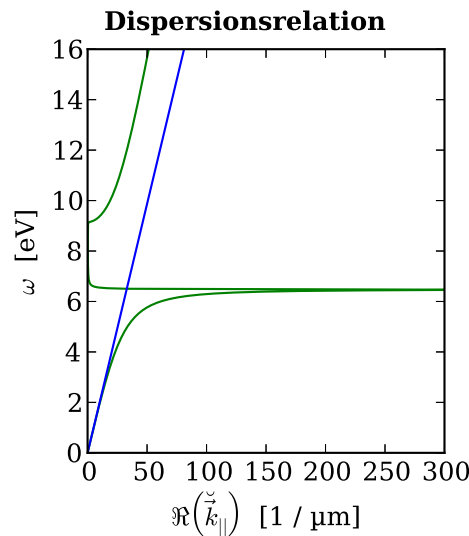
$$\check{k}_{\perp,i}^2 = \omega^2 \check{\mu} \left( \check{\epsilon}_i - \frac{\check{\epsilon}_1 \check{\epsilon}_2}{(\check{\epsilon}_1 + \check{\epsilon}_2)} \right) \quad i \in 1, 2 \quad (2.4.3)$$



**Abbildung 2.4.1.:** Dispersionsrelation des Oberflächenplasmons. Dargestellt sind die Real- und Imaginärteile der parallelen und senkrechten  $\vec{k}$ -Vektoren für eine Luft-Silber-Grenzschicht. Zusätzlich ist in blau die Relation für Luft eingezeichnet. Man beachte, dass die Relation der senkrechten Komponente  $\vec{k}_\perp$  für Luft und Silber im Bereich der Plasmonplasmafrequenz (190 nm) einen ähnlichen Verlauf nimmt.

Für eine Luft-DRUDE-Silber-Grenzschicht, sind die Werte in 2.4.1 aufgetragen<sup>24</sup>. Dabei wird deutlich, dass der Verlauf des Realteils der parallelen Komponente  $\Re(\vec{k}_\parallel)$  dem Verlauf der imaginären Komponente der senkrechten Komponente  $\Im(\vec{k}_{\perp,i})$  ähnelt. Da  $\Im(\vec{k}_{\perp,i})$  ein Maß für die Abklinglänge des Feldes senkrecht zur Oberfläche ist, ist ein Oberflächen-

<sup>24</sup>Die üblichere Darstellung der Dispersionsrelation trägt die Frequenz  $\omega$  als Maß für die Energie über den Realteil der parallelen Komponente  $\Re(\vec{k}_\parallel)$  auf. Diese ist:



plasmon um so stärker an der Oberfläche gebunden, je kürzer seine Wellenlänge  $\lambda_{SPP}$  ist. Darüber hinaus hat das Plasmon im Bereich  $\Im(\check{k}_{\parallel}) < \Re(\check{k}_{\parallel})$ , das heißt, wenn es stärker propagiert, als es absorbiert wird, einen evaneszenten Charakter als einen strahlenden. Die Frequenz, bei der die Plasmonwellenlänge am kürzesten ist, das heißt, wenn  $\Re(\check{k}_{\parallel})$  maximal ist, nennt man auch Plasmafrequenz des Plasmons  $\omega_{SPP}$ . Sie liegt hier bei ca. 190 nm. Der Bereich zwischen der Plasmafrequenz des Silbers 136 nm und der Plasmafrequenz des Plasmons ändert sich die Natur der Eigenlösung dramatisch. So ist nun  $\check{k}_{\parallel}$  im wesentlichen imaginär und  $\check{k}_{\perp}$  reell, das heißt, dass die Lösung auf kurzer Distanz (gemessen an  $\lambda_{SPP}$ ) strahlend zerfällt.

Da im Vergleich mit der „Licht-Linie“ ( $|\vec{k}| = \omega/c_0 < \Re(\check{k}_{\parallel})$ ,  $\omega < \omega_{SPP}$ ) die  $\check{k}_{\parallel}$ -Komponente immer größer ist als der Betrag des  $|\vec{k}|$ -Vektor des Lichtes, das heißt, es existiert keine ebene Welle, welche über die gleiche Translationssymmetrie<sup>25</sup> verfügt, kann diese Eigenlösung der Grenzfläche nicht ohne weitere Konstruktionen durch eine einfallende ebene Welle angeregt werden. Dazu nutzt man zum Beispiel die Anregung mittels Gitter [58] oder totaler interner Reflexion an einer Grenzfläche [30, 52].

Der größere  $k$ -Vektor des Plasmons äußert sich nach Gleichung 2.4.2 gleichzeitig in einer im Vergleich mit der Vakuumlichtwellenlänge teilweise deutlich kleineren Wellenlänge des Plasmons.

---

<sup>25</sup>Die Erhaltung der Translationssymmetrie, beziehungsweise  $\check{k}$ -Erhaltung entspricht der Impulserhaltung für Photonen.

### 2.4.2. Partikel

Neben diesen Oberflächen-Plasmon-Polaritonen kann man auch stark lokalisierte elektromagnetische Felder in metallischen Nanopartikeln als Folge einer kollektiven Elektronenschwingung anregen<sup>26</sup>.

Die Wechselwirkung einer ebenen Welle mit einem sphärischen Partikel wurde von GUSTAV MIE bereits 1908 untersucht [42]. Da die Lösung recht umfangreich ist, soll hier nur der Weg skizziert werden. Eine ausführliche Behandlung findet sich in [4, 42]. Der Kerngedanke besteht darin, die Anregungsfunktion, in diesem Fall die ebene Welle, als Funktion der Besselfunktionen nach Abschnitt 2.1.4 zu entwickeln, um dann die Vollständigkeit und Orthogonalität der Basisfunktionen zu nutzen, um die Stetigkeitsbedingungen (Abschnitt 2.1.3) auf dem Rand der Kugel zu erfüllen.

Bei der Wechselwirkung von Licht mit Metallnanopartikeln (siehe Abbildung 2.4.2) treten verschiedene Phänomene auf. Diese lassen sich in Nah- und Fernfeld unterteilen. Zu den Fernfeldphänomenen gehört die vom Nanopartikel absorbierte beziehungsweise aus dem Lichtweg gestreute Energie. Die abgestrahlte ( $P_S$ ) beziehungsweise absorbierte ( $P_A$ ) Leistung berechnet sich aus dem Integral über die Partikeloberfläche der zeitgemittelten Energieflüsse<sup>27 28</sup>:

$$P_A = \int_{\partial MNP} \langle \vec{S}_A \rangle_t \cdot \vec{n} = \int_{\partial MNP} \frac{1}{2} \Re \left( \vec{\check{E}}_A \times \vec{\check{H}}_A^* \right) \cdot \vec{n} \quad (2.4.4a)$$

$$P_S = \int_{\partial MNP} \langle \vec{S}_S \rangle_t \cdot \vec{n} = \int_{\partial MNP} \frac{1}{2} \Re \left( \vec{\check{E}}_S \times \vec{\check{H}}_S^* \right) \cdot \vec{n} \quad (2.4.4b)$$

Teilt man diese noch durch die mittlere auf die Partikelquerschnittsfläche  $A$  eingestrahlte

<sup>26</sup>Da diese am Partikel gebunden sind, und bei sehr kleinen Partikeln die Form von stehenden Wellen annehmen, werden sie häufig als Lokalisierte-Oberflächen-Plasmonen (LSP von engl.: „Localized Surface Plasmon“) bezeichnet. Diese Bezeichnung ist insofern irreführend, da es sich einerseits um zwei gegensätzlich laufende Plasmonen handelt (wie man sich anhand einer Betrachtung in der Zeitdomäne leicht überzeugen kann) und andererseits der Übergang vom LSP zum laufendem Oberflächenplasmon (PSP von engl.: „Propagating Surface Plasmon“) an einem Nanodraht fließend ist. Aus diesem Grund wird in dieser Arbeit nicht zwischen den verschiedenen Plasmonarten unterschieden.

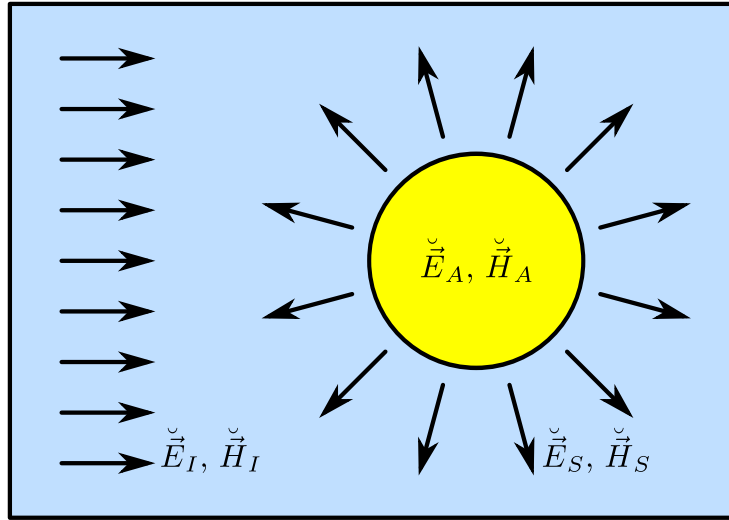
<sup>27</sup>Der Energiefluss eines elektromagnetischen Feldes wird durch den Poyntingvektor:

$$\vec{S} = \vec{E} \times \vec{H} = \Re \left( \vec{\check{E}} \right) \times \Re \left( \vec{\check{H}} \right)$$

angegeben. Der Energiefluss im Zeitmittel  $\langle \vec{S} \rangle_t$  errechnet sich aus:

$$\langle \vec{S} \rangle_t = \frac{1}{2} \Re \left( \vec{\check{E}} \times \vec{\check{H}}^* \right)$$

<sup>28</sup>Hier ist  $\vec{n}$  der Normalenvektor der Integrationsfläche



**Abbildung 2.4.2.:** Streuung- und Absorption an einem Nanopartikel. Für die Berechnung wird das Feld in drei Komponenten zerlegt: das einfallende Feld ( $\vec{E}_I, \vec{H}_I$ ), das gestreute Feld ( $\vec{E}_S, \vec{H}_S$ ) und das Feld im Partikel ( $\vec{E}_A, \vec{H}_A$ ).

Intensität  $I_I$ , so erhält man den Streu- ( $C_S$ ) beziehungsweise Absorptionsquerschnitt ( $C_A$ ):

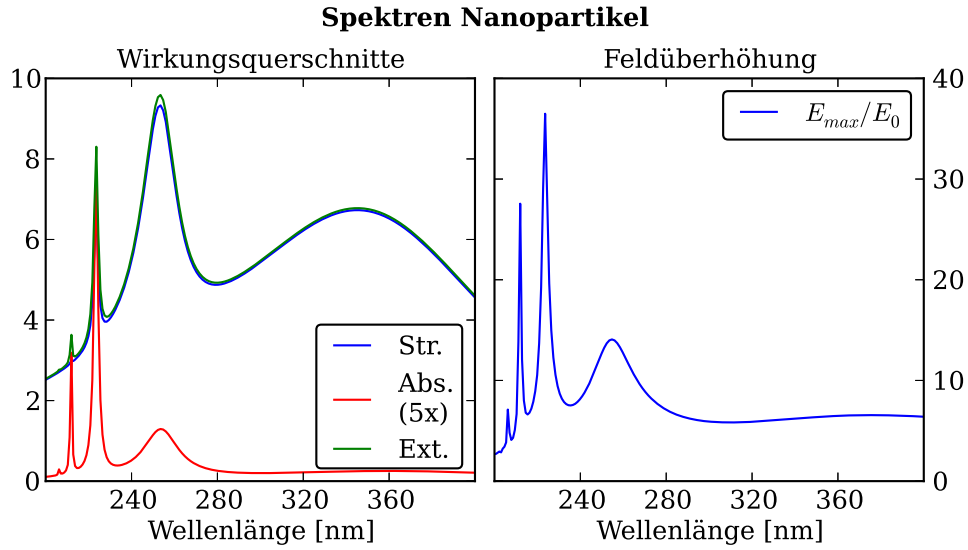
$$C_S = P_S/I_I \qquad C_A = P_A/I_I \qquad (2.4.5)$$

Will man die gesamte aus dem Lichtstrahl entfernte, das heißt, entweder absorbierte oder herausgestreute, Energie wissen, so nutzt man dazu den Extinktionsquerschnitt  $C_E$ :

$$C_E = C_A + C_S \qquad (2.4.6)$$

Diese Querschnitte werden im Folgenden, da sie die Antwort eines Partikels auf eine äußere Anregung abbilden, als Wirkungsquerschnitte bezeichnet. Da sich diese Werte im Allgemeinen schlecht vorstellen lassen, wird in dieser Arbeit, soweit nicht anders erwähnt, immer mit den normierten, das heißt auf die Partikelquerschnittsfläche bezogenen, Querschnitten ( $C := C/A$ ) gerechnet.

Neben diesen Fernfeldphänomenen ist speziell im Bereich der Plasmonik das Nahfeld von besonderem Interesse. Die Ursache liegt in den eingangs erwähnten kollektiven Elektronenschwingungen. Diese sorgen für eine Feldüberhöhung in der Nähe der Partikeloberfläche. Beide Phänomene sind in Abbildung 2.4.3 als Funktion der Wellenlänge dargestellt. Deutlich sind die einzelnen Maxima in den Spektren zu sehen. Diese kann man nach dem dominierenden Term in der Multipolentwicklung (siehe [4, 14]) klassifizieren. Während sich



**Abbildung 2.4.3.:** Wirkungsspektren und Nahfeldüberhöhung an einem 100 nm großen Silber-Nanopartikel. Die Absorption wurde um den Faktor fünf überhöht dargestellt.

die Dipolresonanz noch in quasistatischer Näherung ermitteln lässt<sup>29</sup>, sind die höheren Resonanzen eine Folge der Retardierung des elektromagnetischen Feldes. Die zugehörigen Feldbilder sind in Abbildung 2.4.4 für die ersten<sup>30</sup> vier Maxima gezeigt. Im Allgemeinen kann man feststellen, dass mit höherer Resonanzordnung, die Streuung ab- und Absorption zunimmt. Des Weiteren ist das Feld mit höherer Resonanzordnung stärker um das Partikel lokalisiert und besitzt eine höhere Güte ( $\omega/\Delta\omega$ ). Diese Eigenschaften findet man nicht nur für runde Nanopartikel. Auch andere (endliche) Systeme zeigen ähnliche Eigenschaften. Diese sind in den seltensten Fällen noch analytisch zugänglich, so dass zunehmend

<sup>29</sup>Für runde Partikel in statischer Näherung ist das Dipolmoment  $\check{p}$  (mit Medium 1 die Umgebung und Medium 2 das Partikel) gegeben durch:

$$\check{p} = \check{\epsilon}_2 \check{\alpha} \check{E}_0 \quad \check{\alpha} = 3V \frac{\check{\epsilon}_2 - \epsilon_1}{\check{\epsilon}_2 + 2\epsilon_1}$$

Hier sind  $V$  das Volumen des Partikels und  $\check{E}_0$  das äußere elektrische Feld am Ort des Partikels. Die Streu- und Absorptionsquerschnitte ermitteln sich dann aus:

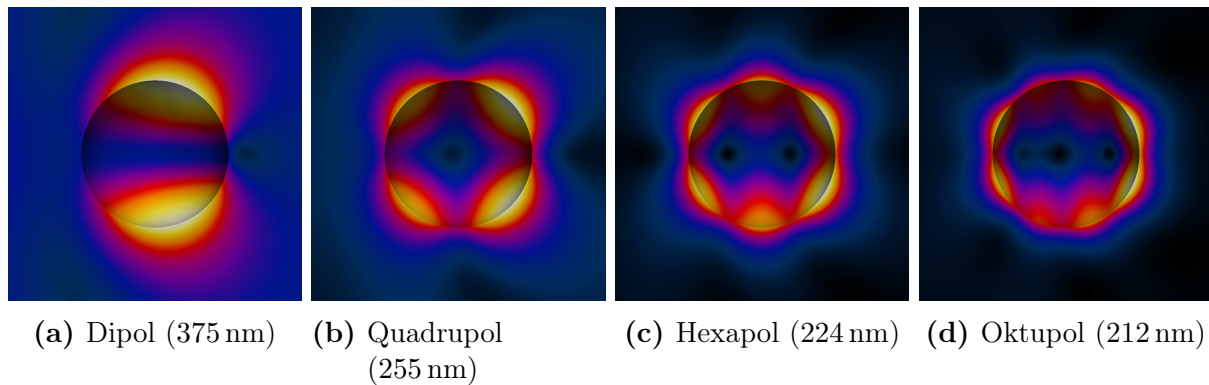
$$C_E = k_1 \Im(\check{\alpha}) \quad C_S = \frac{k_1^4}{6\pi} |\check{\alpha}|^2$$

Das heißt, die Resonanzbedingung für den Dipol hängt demnach nur von der Polarisierbarkeit  $\check{\alpha}$  ab und lautet:

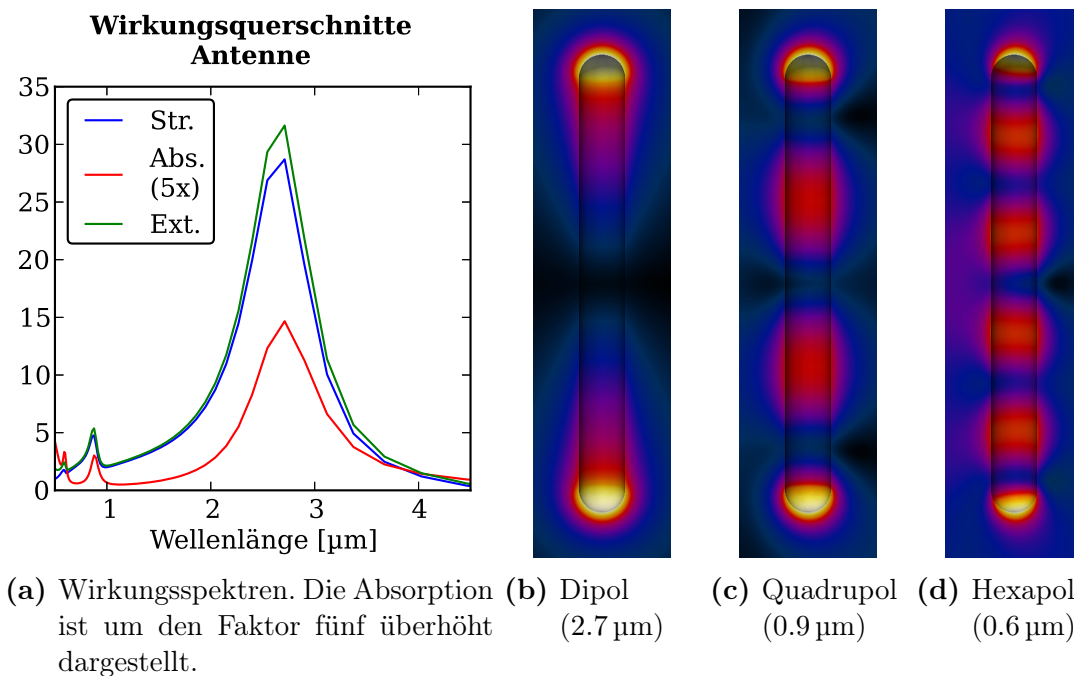
$$\check{\epsilon}_2 + 2\epsilon_1 \rightarrow \min$$

<sup>30</sup>Dabei beginnt die Zählung von kleinen Energien und damit von großen Wellenlängen.





**Abbildung 2.4.4.:** Felddbilder an einem 100 nm großen Silber-Nanopartikel. Die gewählten Wellenlängen entsprechen den ersten (langwelligen) Resonanzen aus Abbildung 2.4.3.



**Abbildung 2.4.5.:** Spektren und Nahfeld für eine  $1 \mu\text{m} \times 100 \text{ nm}$  große Gold-Antenne.

numerische Methoden zum Einsatz kommen. Ein weiteres Beispiel zeigt Abbildung 2.4.5, in dem die Wirkungsspektren samt zugehöriger Felddbilder an den Resonanzpositionen für eine  $1 \mu\text{m} \times 100 \text{ nm}$  große Gold-Antenne gezeigt sind.

Die Resonanzpositionen und Höhen sind im Allgemeinen vom Material und der Form abhängig. Dies eröffnet die Möglichkeit, durch geeignete Strukturierungsmaßnahmen diese gezielter einzustellen.

## 2.5. Frequenzverdopplung in Metallen

Bei der Erzeugung der zweiten Harmonischen<sup>31</sup> handelt es sich um einen nichtlinearen Prozess in Medien, bei dem durch Wechselwirkung mit Licht der Frequenz  $\omega$ , Strahlung der doppelten Frequenz  $2\omega$  entsteht. Da die Energie eines Photons, mit der Frequenz gekoppelt ist, bedeutet dies, dass ein Energieübertrag von Licht der Frequenz  $\omega$  nach Licht der Frequenz  $2\omega$  stattfindet. In linearen lokalen isotropen Materialien gilt 2.1.10:

$$\vec{P}(r, t) = \int dt' \chi(r, t, t') \vec{E}(r, t') \quad \check{\vec{P}}(r, \omega) = \check{\chi}(r, \omega) \check{\vec{E}}(r, \omega)$$

$\check{\vec{P}}$  und damit auch  $\partial_t \check{\vec{P}}$  sind hier linear  $\check{\vec{E}}$  abhängig, damit ist keine Frequenzverdopplung möglich. Um dieses Phänomen dennoch beschreiben zu können wird, die Abhängigkeit von  $\vec{P}$  als TAYLOR-Entwicklung von  $\vec{E}$  aufgeschrieben (mit  $\vec{E}(r, t) = \int d\vec{k} d\omega \check{\vec{E}}(\vec{k}, \omega) e^{i(\vec{k}\vec{r} - \omega t)}$  [43]):

$$\begin{aligned} P_i(\vec{r}, t) &= P_i^{(1)}(\vec{r}, t) + P_i^{(2)}(\vec{r}, t) + \dots \\ &= \epsilon_0 \sum_j \int d\vec{r}' dt' \chi_{ij}^{(1)}(\vec{r} - \vec{r}', t - t') E_j(\vec{r}', t') \\ &\quad + \epsilon_0 \sum_{jk} \int d\vec{r}' d\vec{r}'' dt' dt'' \chi_{ijk}^{(2)}(\vec{r} - \vec{r}', \vec{r} - \vec{r}'', t - t', t - t'') E_j(\vec{r}', t') E_k(\vec{r}'', t'') \\ &\quad + \dots \\ &= \underbrace{\epsilon_0 \int d\vec{k}' d\omega' e^{i\vec{k}'\vec{r}} e^{-i\omega' t} \sum_j \check{\chi}_{ij}^{(1)}(\vec{k}', \omega') \check{E}_j(\vec{k}', \omega')}_{\chi_{ij} E_j} \\ &\quad + \underbrace{\epsilon_0 \int d\vec{k}' d\vec{k}'' d\omega' d\omega'' e^{i(\vec{k}' + \vec{k}'')\vec{r}} e^{-i(\omega' + \omega'')t} \sum_{jk} \check{\chi}_{ijk}^{(2)}(\vec{k}', \vec{k}'', \omega', \omega'') \check{E}_j(\vec{k}', \omega') \check{E}_k(\vec{k}'', \omega'')}_{\chi_{ijk} E_j E_k} \\ &\quad + \dots \end{aligned} \tag{2.5.1}$$

Die Verallgemeinerung auf beliebige Ordnungen lautet dann:

$$P_i = \chi_{ij}^{(1)} E_j + \chi_{ijk}^{(2)} E_j E_k + \chi_{ijkl}^{(3)} E_j E_k E_l + \mathcal{O}(E^4) \tag{2.5.2}$$

<sup>31</sup>engl.: „Second Harmonic Generation“ (SHG)

Betrachtet man nun  $P^{(2)}$  mit  $E_j, E_k$  der Form  $E = \Re(e^{-i\omega_j}) = \frac{1}{2}(e^{-i\omega_j} + e^{i\omega_j})$ , dann enthält  $P^{(2)}$  folgende Komponenten<sup>32</sup>:

- $2\omega_j, 2\omega_k$  Frequenzverdopplung
- $\omega_j + \omega_k$  Summenfrequenz
- $\omega_j - \omega_k$  Differenzfrequenz
- $0$  optische Gleichrichtung

Für den Fall, dass  $\omega_j = \omega_k$  ist, führt die Summenfrequenz ebenfalls zur Frequenzverdopplung. Nun stellt sich die Frage nach der Größe von  $\chi_{ijk}$ . Dazu wird folgende Inversionssymmetrie betrachtet: (aus  $E \rightarrow -E$  folgt  $P \rightarrow -P$ ), d.h.:  $\sum_{jk} \chi_{ijk} E_j E_k = -\sum_{jk} \chi_{ijk} (-E_j)(-E_k)$ . Daraus folgt, dass  $\chi_{ijk} \equiv 0$  sein muss.

### Woher stammt die SHG von Metallen?

Der obige Zusammenhang stellt im Wesentlichen das Verhalten innerhalb von Festkörpern dar. An Oberflächen gilt die Aussage mit der Inversionssymmetrie nicht mehr, da die Oberfläche diese Symmetrie verletzt. So klingt innerhalb von  $5 - 10 \text{ \AA}$  (siehe [33, 34, 68]) die Elektronendichte des Metalls ab. Auch wurden in obiger Herleitung Einflüsse wie die LORENTZ-Kraft vernachlässigt. Eine allgemeine Herleitung zur Lösung der MAXWELL-Gleichungen an nichtlinearen Oberflächen findet sich in [3].

Für Metalloberflächen wurde in [19] mit Hilfe des Ansatzes der „quasifreien“ Elektronen folgender Zusammenhang hergeleitet:

$$\check{P}^{(2\omega)} = \check{\alpha}_1 \check{E}^{(1\omega)} (\vec{\nabla} \cdot \check{E}^{(1\omega)}) + \check{\alpha}_2 \vec{\nabla} (\check{E}^{(1\omega)} \cdot \check{E}^{(1\omega)}) \quad (2.5.3)$$

Dabei beschreibt der erste Term den Oberflächenanteil, welcher die Dichteschwankung der Elektronen an der Metalloberfläche berücksichtigt, und der zweite Term den Volumenanteil, welcher im Wesentlichen durch die LORENTZ-Kraft hervorgerufen wird. In [8, 59] wurden diese Zusammenhänge durch Quantenmechanische Rechnungen verifiziert und verfeinert. In [61] wurde das Modell in der hier verwendeten Form inklusive des THOMAS-FERMI-Druckes verwendet und der Zusammenhang mit Gleichung 2.5.3 dargestellt. Eine kurze Herleitung dazu findet sich in [57] und soll hier kurz wiedergegeben werden.

Begonnen wird mit dem Hydrodynamischen Modell nach Gleichung 2.3.21 :

$$\begin{aligned} \partial_t n^- + \nabla \cdot (n^- \vec{v}) &= 0 \\ \partial_t (n^- \vec{v}) + \nabla \cdot \left( n^- \vec{v} \otimes \vec{v} + \frac{1}{m_e} p \right) &= -\frac{e_0}{m_e} n^- \left( \vec{E} + \mu \vec{v} \times \vec{H} \right) - \frac{1}{\tau} n^- \vec{v} \end{aligned}$$

<sup>32</sup>Sei hier ohne Beschränkung der Allgemeinheit  $\omega_j \geq \omega_k$ .

Nun kann man zum Frequenzraum übergehen und die wesentlichen Variablen<sup>33</sup> ( $\rho = n^-$  und  $\vec{v}$ ) nach linearen und nichtlinearen Anteilen entwickeln:

$$\check{\rho} = \rho^{(0)} + \check{\rho}^{(1)} e^{-i\omega t} + \check{\rho}^{(2)} e^{-i2\omega t} + \dots \quad (2.5.4a)$$

$$\check{\vec{v}} = \check{\vec{v}}^{(1)} e^{-i\omega t} + \check{\vec{v}}^{(2)} e^{-i2\omega t} + \dots \quad (2.5.4b)$$

Nimmt man an, dass das lineare Feld durch die Erzeugung der SHG weitgehend unbeeinflusst bleibt, so ergeben sich folgende Zusammenhänge<sup>34</sup>:

$$\partial_t \check{\rho}^{(1)} = -i\omega \rho^{(1)} = \rho^{(0)} \vec{\nabla} \check{\vec{v}}^{(1)} \quad (2.5.5a)$$

$$\partial_t \check{\vec{v}}^{(1)} = -i\omega \check{\vec{v}}^{(1)} = -\frac{e_0}{m} \check{\vec{E}}^{(1)} \quad (2.5.5b)$$

$$\partial_t \check{\vec{v}}^{(2)} = -i2\omega \check{\vec{v}}^{(2)} = -(\check{\vec{v}}^{(1)} \cdot \vec{\nabla}) \check{\vec{v}}^{(1)} - \frac{e_0}{m} \mu(\check{\vec{v}}^{(1)} \times \check{\vec{H}}^{(1)}) \quad (2.5.5c)$$

Mit Hilfe der Zusammenhänge:

$$\partial_t \check{\vec{P}}^{(1)} = -i\omega \check{\vec{P}}^{(1)} = \check{\vec{j}}^{(1)} \quad \partial_t \check{\vec{P}}^{(2)} = -i2\omega \check{\vec{P}}^{(2)} = \check{\vec{j}}^{(2)} \quad (2.5.6a)$$

$$\check{\vec{j}}^{(1)} = \rho^{(0)} \check{\vec{v}}^{(1)} \quad \check{\vec{j}}^{(2)} = \check{\rho}^{(1)} \check{\vec{v}}^{(1)} + \rho^{(0)} \check{\vec{v}}^{(2)} \quad (2.5.6b)$$

und  $\vec{\nabla} \times \check{\vec{E}} = i\omega \mu \check{\vec{H}}$  lassen sich die lineare und nichtlineare Polarisation abschätzen zu:

$$\check{\vec{P}}^{(1)} = \frac{\rho^{(0)} e_0}{m_e \omega^2} \check{\vec{E}}^{(1)} \quad (2.5.7a)$$

$$\check{\vec{P}}^{(2)} = -\frac{\rho^{(0)} e_0^2}{m_e^2 \omega^4} \left[ \check{\vec{E}}^{(1)} (\vec{\nabla} \cdot \check{\vec{E}}^{(1)}) + \frac{1}{4} \vec{\nabla} (\check{\vec{E}}^{(1)} \cdot \check{\vec{E}}^{(1)}) \right] \quad (2.5.7b)$$

Bei der Herleitung wurden die Dämpfungs- und Druckterme vernachlässigt. Bemerkenswert ist hier, dass Beiträge der Form  $\vec{\nabla} \cdot \check{\vec{E}}$  auftauchen, wo doch  $\vec{\nabla} \cdot \vec{E} = 0$  gelten soll. Dieser Term ist demnach nur an Oberflächen von Bedeutung und beschreibt im Wesent-

<sup>33</sup>Zur besseren Übersichtlichkeit, und da ohnehin nur die negativen Ladungsträger einen Beitrag liefern, wird im Folgenden die Teilchendichte als  $\rho = n^-$  bezeichnet.

<sup>34</sup>Bei der letzten Gleichung wurde ausgenutzt, dass:

$$\begin{aligned} \partial_t(\rho \vec{v}) &= (\partial_t \rho) \vec{v} + \rho (\partial_t \vec{v}) \\ \vec{\nabla} \cdot (\rho \vec{v} \otimes \vec{v}) &= (\vec{\nabla} \cdot \vec{v}) \rho \vec{v} + \vec{v} \cdot (\vec{\nabla} \rho \vec{v}) \end{aligned}$$

Damit wird Gleichung 2.3.21b:

$$\partial_t(\vec{v}) + (\vec{v} \cdot \vec{\nabla}) \vec{v} + \frac{1}{\rho m_e} \vec{\nabla} p = -\frac{e_0}{m_e} \left( \vec{E} + \mu \vec{v} \times \vec{H} \right) - \frac{1}{\tau} \vec{v}$$

lichen die Elektronendichteschwankung. Die Frage, ob nun der Oberflächenbeitrag oder der Volumenbeitrag für die SHG von Metallen wichtiger ist, lässt sich nicht so einfach beantworten, es gibt jedoch Tendenzen, die darauf hindeuten, dass die SHG auf Grund der Elektronendichteschwankung zusammen mit dem hohen Feld an der Metalloberfläche die SHG dominiert. Da in dieser Arbeit immer mit dem vollen Modell gearbeitet wird (siehe Kapitel 5.3), soll im Folgenden nicht weiter zwischen beiden Beiträgen unterschieden werden.

## 2.6. Numerische Verfahren

Neben den theoretischen Grundlagen aus den vorangegangenen Abschnitten und den Details der DG-Methode in Kapitel 3 ist für das Verständnis der Numerik weiteres Handwerkzeug notwendig, welches in diesem Abschnitt erläutert werden soll. Dazu gehört einerseits die Begründung für die Wahl des numerischen Verfahrens, wie auch der exemplarische Aufbau einer Simulation.

### 2.6.1. Einteilung der numerischen Verfahren

Wenn es an die numerische Behandlung der MAXWELL-Gleichung geht, stehen verschiedene Methoden zur Auswahl. Diese lassen sich einerseits nach der Art des kausalen Zusammenhangs (Frequenz- oder Zeitdomäne) und nach der Art der Diskretisierung, das heißt, der Zerlegung des zu behandelnden Problems unterteilen (Rand- oder Volumenverfahren).

#### Frequenz- vs. Zeitdomäne

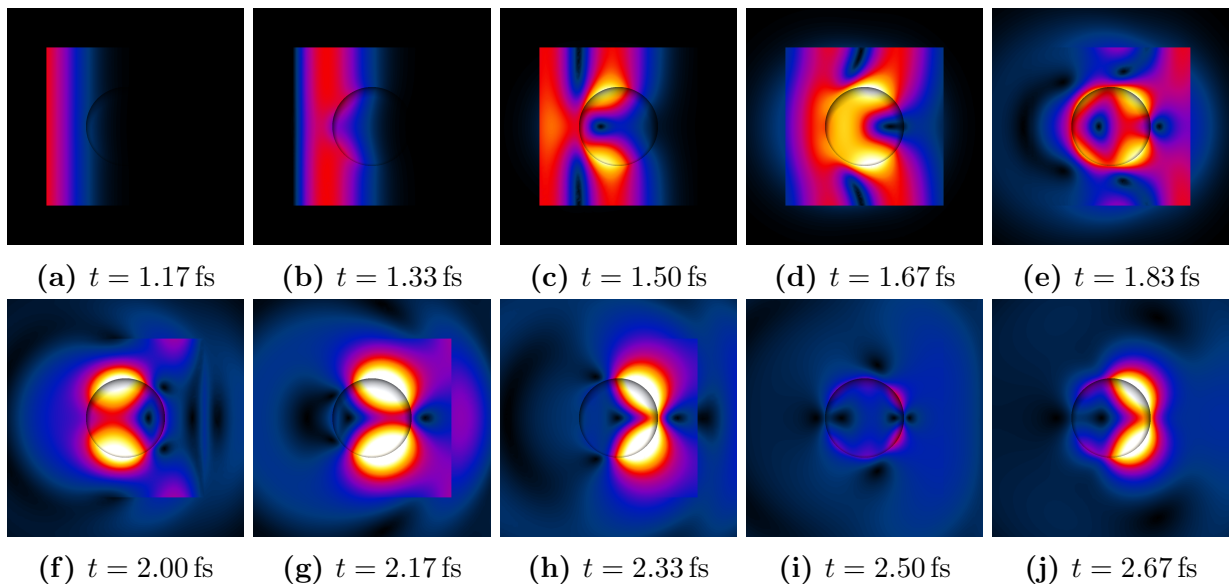
Bei Verfahren in der Frequenzdomäne wird analog zur Lösung der MAXWELL-Gleichung in Abschnitt 2.1.4 die Lösung für eine spezielle Frequenz  $\omega$  gesucht. Dabei hat man den Vorteil, dass die Zeitableitungen via Ansatz bereits absepariert sind und man nur noch die Ortsableitungen der Gleichungen behandeln muss. Das heißt aber auch, dass der zeitliche Verlauf der Felder, insbesondere die Einschwingvorgänge und die Gruppendynamik von Wellenpaketen (siehe Abbildung 2.6.1, 2.6.5 und 2.6.6 in den nachfolgenden Abschnitten) nicht mit berücksichtigt wird. Als Vorteil ist in diesem Zusammenhang die Behandlung periodischer Strukturen zu nennen, bei denen nun nur eine Elementarzelle berücksichtigt werden muss. Die Wechselwirkung mit den Nachbarzellen lässt sich leicht über Phasenfaktoren an den Elementarzellengrenzen berücksichtigen. Da diese Verfahren häufig nur eine Frequenz berücksichtigen sind nichtlineare Materialmodelle, bei denen es zu einer Wechselwirkung verschiedener Frequenzkomponenten des Lichtes kommt, nur schwer zu modellieren. Sie werden daher häufig nur im Rahmen störungstheoretischer Ansätze behandelt<sup>35</sup>.

Demgegenüber stehen die Zeitschrittverfahren, bei denen die zeitabhängigen MAXWELL-Gleichungen (2.1.1) beginnend von einem bekannten Zustand des Feldes zum Zeitpunkt  $t_0$  Zeitschritt für Zeitschritt berechnet werden (siehe Abbildung 2.6.1). Bei denen ist, ein geeignetes Materialmodell vorausgesetzt, auch die Wechselwirkung von verschiedenen Frequenzkomponenten des Lichtes möglich. Auch lassen sich hier die zeitlichen Variationen der

<sup>35</sup>So kann man wie im Abschnitt 2.5 erst die Lösung bei der Fundamentalfrequenz  $\omega$  berechnen und die entstehende SHG (bei  $2\omega$ ) im Rahmen einer Abschätzung ermitteln.

Felder (z.B.: die Ausbreitung eines Pulses beziehungsweise die Physik von Einschwingvorgängen) gut studieren. Nachteilig ist, dass um die Lösung für eine spezielle Wellenlänge zu bekommen immer der komplette Zeitverlauf berechnet und FOURIER-transformiert werden muss. Auch lassen sich periodische Systeme, ohne weitere Maßnahmen, nur approximativ lösen, das heißt, dass das Gitter durch eine endliche Anzahl an Elementarzellen genähert wird. Nachteilig ist ebenso, dass bei Systemen mit sehr langen Einschwingzeiten, auch sehr lange gerechnet werden muss, bis man eine ausreichend genaue Lösung erhält. Demgegenüber erhält man nach der FOURIER-Transformation aber auch gleich das ganze Spektrum und nicht nur die Werte für einen einzigen Frequenzpunkt.

Typische Vertreter für Frequenzdomänmethoden sind MMP (siehe [14]) und das klassische FEM (siehe [21]). Vertreter für Zeitschrittverfahren sind FDTD (siehe [64, 72]) und die in dieser Arbeit verwendete Variante von DG<sup>36</sup>.



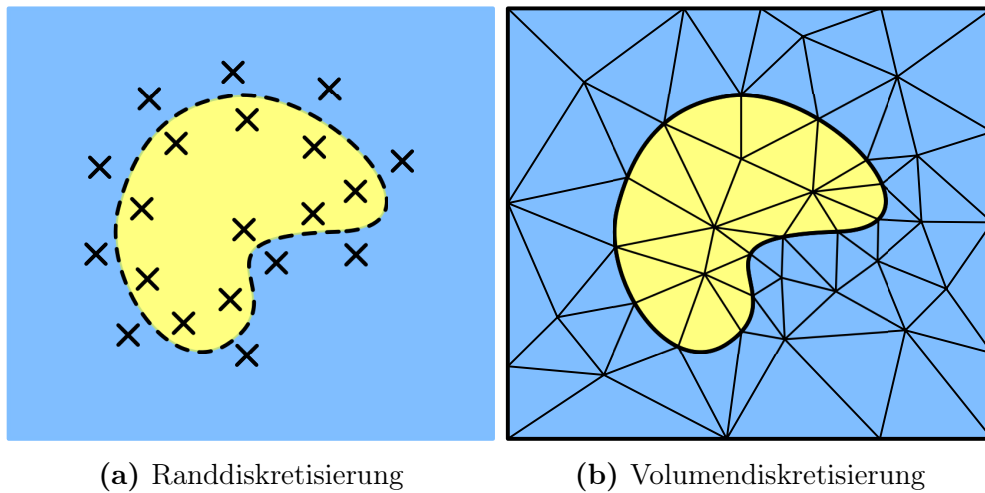
**Abbildung 2.6.1.:** Zeitevolution des  $E$ -Feldes eines 100 nm Silber-Partikels. Vergleiche Kapitel ?? und Abbildung 2.4.3 und 2.4.4

### Rand- vs. Volumendiskretisierung

Damit ein System vom Computer berechnet werden kann, muss es geeignet zerlegt werden. Dies soll hier am Beispiel der Wechselwirkung des Lichtes mit einem Partikel skizziert werden (siehe Abbildung 2.6.2).

Die randdiskretisierenden Verfahren unterteilen dabei lediglich den Rand des Partikels und beschreiben die Wechselwirkung mit Hilfe von Hilfsquellen wie Oberflächenladungen

<sup>36</sup>Mittlerweile sind auch FD-Implementationen von DG bekannt [7]



**Abbildung 2.6.2.:** Vergleich zwischen Rand- und Volumendiskretisierung

und -Strömen oder Multipolen (Abbildung 2.6.2a) (siehe [10, 14, 24]). Dies hat den Vorteil, dass die MAXWELL-Gleichungen nur auf dem Rand des Partikels gelöst werden müssen. Das führt zu einem verhältnismäßig kompakten Gleichungssystem. Nachteilig ist, dass die einzelnen Quellterme der Felder an jedem Punkt der Oberfläche einen nicht verschwindenden Beitrag liefern. Daher hängen die Koeffizienten der Quellterme (über die Oberfläche) mit allen anderen Quelltermen zusammen, das heißt, das entstehende Gleichungssystem ist dicht besetzt<sup>37</sup>. Dies führt numerisch insbesondere bei komplizierten Systemen zu Komplikationen (siehe [62]).

Demgegenüber stehen die volumendiskretisierenden Verfahren. Diese unterteilen das ganze Volumen (Abbildung 2.6.2b) von Partikel und Umgebung in einzelne Stücke und stellen die Felder mit Hilfe von Basisfunktionen auf diesen Elementen dar. Da nun nur benachbarte Elemente (beziehungsweise deren Felder) wechselwirken, das heißt, die Feldkoeffizienten nur von den Feldern der näheren Umgebung abhängen, nennt man das entstehende Gleichungssystem dünn besetzt. Nachteilig ist, dass nun deutlich mehr Unbekannte berechnet werden müssen, da nicht nur die Oberfläche sondern das Volumen unterteilt wird und da nicht nur das Volumen des Partikels, sondern auch das des umgebenden Gebietes behandelt werden muss. Dies führt unweigerlich zu der Frage, wann und wie man dieses Gebiet abbricht, ohne die Lösung zu stark negativ zu beeinflussen.

Typische Vertreter der randdiskretisierenden Methoden sind MMP und die Randintegralmethoden ([10, 14, 24]). Vertreter für volumendiskretisierende Methoden sind FEM, DG und FDTD<sup>38</sup>.

<sup>37</sup>Die Bezeichnung dicht beziehungsweise dünn besetzt rührt aus der Betrachtung der Einträge der Matrix des Gleichungssystems

<sup>38</sup>Bei der FDTD-Methode ist die Unterteilung, zumindest in der bekanntesten Ausprägung nicht in Drei-



## Warum DG?

Trotz der scheinbaren Nachteile, die volumendiskretisierende Verfahren mit sich bringen, überzeugen sie vor allem durch leichte Anwendbarkeit, Allgemeinheit und Flexibilität. Insbesondere zeigen sie sich bei komplexen Systemen gegenüber denen der randdiskretisierenden Verfahren im Vorteil ([62]). Daher soll auch hier der Fokus auf diesem Verfahrenstyp liegen. Des Weiteren sollen in dieser Arbeit auch nichtlineare Effekte, insbesondere das spektrale Verhalten der SHG (siehe Abschnitt 5.3) mit Hilfe von A priori Modellen untersucht werden, so dass hier ein Zeitschrittverfahren zur Anwendung kommen soll. Von den bekanntesten Volumenverfahren (FDTD und FEM) hat FEM, auf Grund der nicht-uniformen Diskretisierung (Zerlegung des Simulationsgebietes in meistens Dreiecke und Tetraeder), den Vorteil der guten räumlichen Auflösung, während sich die FDTD-Methode durch ein explizites Zeitschrittverfahren gut für nichtlineare Rechnungen eignet. Die Kernidee der DG-Methode ist nun beide Vorteile miteinander zu kombinieren.

### 2.6.2. Aufbau bei DG

Der generelle Aufbau einer DG-Simulation ist in Abbildung 2.6.3 an Hand der Streuung des Lichtes an einem Partikel, welches sich über einem Substrat befindet, dargestellt. Dabei wird das Simulationsgebiet in vier Kernbereiche unterteilt:

- Gebiet 1 der bzw. die Streuer
- Gebiet 2 die innere Umgebung des Streuers
- Gebiet 3 die äußere Umgebung des Streuers
- Gebiet 4 der PML

Die Gebiete 1 & 2 werden zum Gesamtfeld<sup>39</sup> zusammengefasst und die Gebiete 3 & 4 zum Streufeld<sup>40</sup>. Die Grenze (in Abbildung 2.6.3 grün dargestellt) ist die TF/SF-Kontur. Diese dient der Injektion der Anregung, welche weiter unten behandelt wird. Gleichzeitig dient sie als Integrationsgrenze zur Ermittlung der Streu- und Absorptionsquerschnitte. Da bei Volumenverfahren das Simulationsgebiet abgeschnitten werden muss, ist es wichtig, dass vom Partikel weggestreute Felder nicht am Rand reflektiert werden, und so wieder in das Simulationsgebiet zurücklaufen können. Dazu bedient man sich einer Antireflexschicht<sup>41</sup> (Gebiet 4), welche zur Aufgabe hat, das gestreute Feld ohne Reflexion zu absorbieren. Die

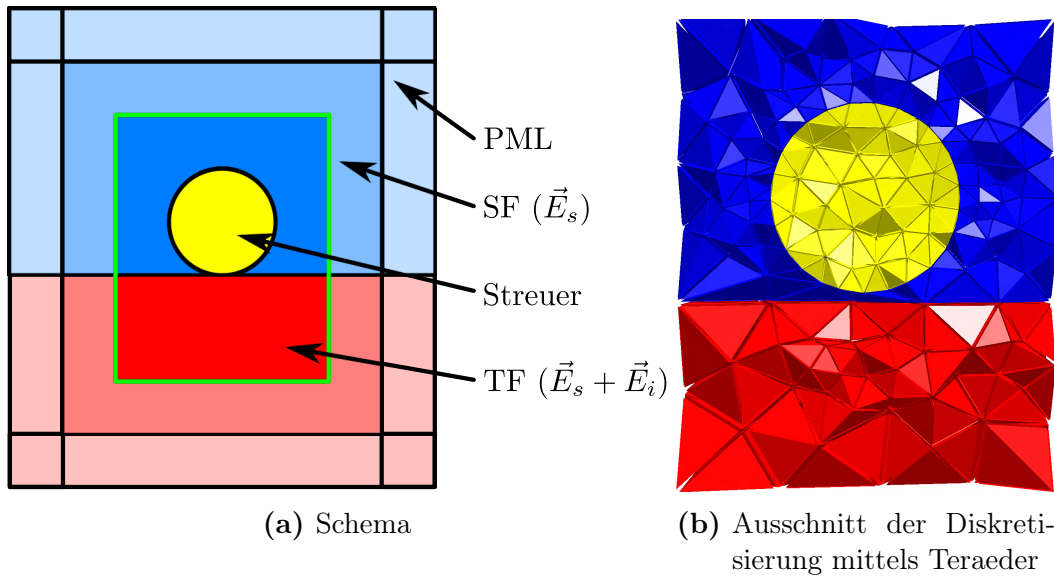
---

ecke beziehungsweise Tetraeder, sondern das YEE-Grid [64, 72], es sind aber auch dreieckige Diskretisierungen bekannt [36].

<sup>39</sup>engl.: „Total Field“ (TF)

<sup>40</sup>engl.: „Scattered Field“ (SF)

<sup>41</sup>engl.: „Perfectly Matched Layer“ (PML)



**Abbildung 2.6.3.:** Schematischer Aufbau einer Simulation. Bei der Konstruktion der PML-Gebiete ist zu bedenken, dass der PML eine richtungsabhängige Absorption zeigt. Aus diesem Grund sind die Kanten und Ecken extra unterteilt.

Konstruktion eines solchen Materials wurde in Abschnitt 2.3.2 gezeigt.

### 2.6.3. Anregungsfelder

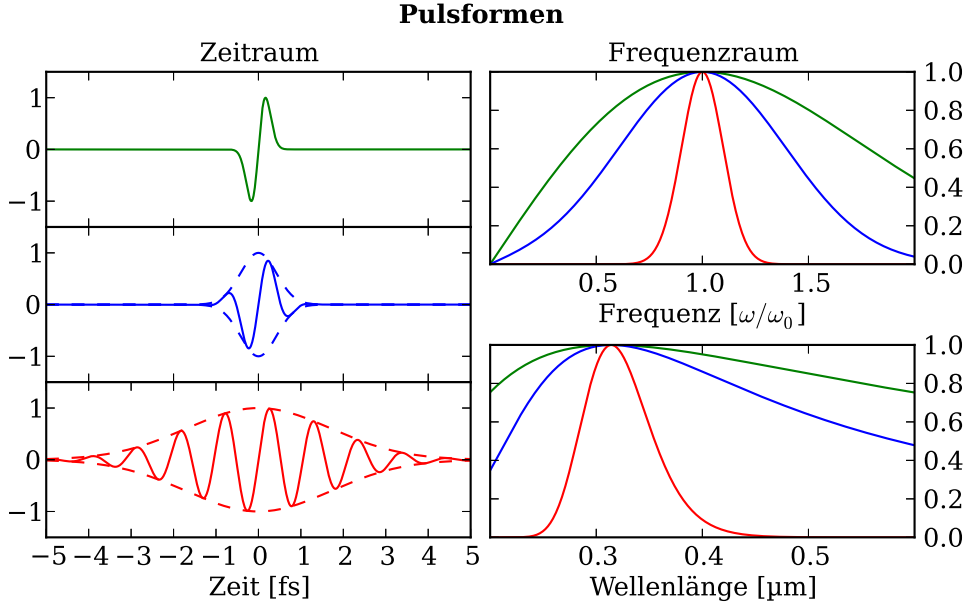
Um nun Quellen in das Simulationsgebiet zu injizieren, gibt es zwei Möglichkeiten. Einerseits wird bereits beim Start eine Lösung vorgegeben. Diese Variante kommt insbesondere bei der Untersuchung von Kavitäten zum Einsatz. Oder man regt das System mit einem äußeren Feld an. Letztere Methode wird häufig bei „Streuexperimenten“ gewählt.

Da bei der DG-Methode nur zwei, die zeitunabhängigen, der vier MAXWELL-Gleichungen verwendet werden, muss bereits zu Beginn der Simulation mit divergenzfreien Feldern begonnen werden. Aus praktischen Gründen wählt man  $\vec{E} = \vec{H} \equiv 0$ . Um nun die Anregung in das Simulationsgebiet einzuschleusen, nutzt man die Linearität der MAXWELL-Gleichungen. Man teilt dabei das Feld auf in:

$$\vec{E}_A = \vec{E}_S + \vec{E}_I \qquad \vec{H}_A = \vec{H}_S + \vec{H}_I \qquad (2.6.1)$$

Dabei bezeichnet  $\vec{E}_I$  bzw.  $\vec{H}_I$  das anregende Feld,  $\vec{E}_A$  bzw.  $\vec{H}_A$  das gesamte Feld und  $\vec{E}_S$  bzw.  $\vec{H}_S$  das gestreute Feld. Diese Einteilung hat zudem den Vorteil, dass man so die Streu- und Absorptionsquerschnitte leicht berechnen kann (Vergleiche dazu Abschnitt 2.4.2). Die Unterteilung selbst wird an der TF/SF-Kontur vorgenommen, das heisst, dass

das einfallende Feld an der Kontur in Richtung Inneres addiert und in Richtung Äußeres subtrahiert wird. Als Anregungsfelder kommen hauptsächlich Pulse zum Einsatz, da sie endlich in der Zeit sind und langsam einschwingen<sup>42</sup>. Für den räumlichen Verlauf der Anregungsfelder eignen sich Multipole und ebene Wellen (siehe Abschnitt 2.1.4). Als zeitlicher Verlauf kommen in dieser Arbeit asymmetrische Pulse mit GAUSSchem beziehungsweise differenziertem GAUSSchem Profil zum Einsatz. Diese Asymmetrie sorgt einerseits für ei-



**Abbildung 2.6.4.:** Auf Eins normierte Pulsformen in Zeit- und Frequenzdomäne. Grün: diff. GAUSS-Puls mit  $\lambda_0 = 314$  nm, Blau: GAUSS-Puls mit  $\lambda_0 = 314$  nm und  $\Delta t = 1/\Delta\omega = 0.4$  fs und Rot: GAUSS-Puls mit  $\lambda_0 = 314$  nm und  $\Delta t = 1/\Delta\omega = 1.7$  fs. Für die GAUSSschen Pulse ist zusätzlich die Einhüllende (gestrichelt) eingezeichnet.

ne Gleichspannungsfreiheit und für ein rein imaginäres Frequenzspektrum. Sie haben die allgemeine Form:

$$E_y = E_0 \mathcal{F}(t) \quad H_z = \frac{E_0}{Z} \mathcal{F}(t) \quad (2.6.2)$$

Mit dem GAUSSchem Profil (Mittenfrequenz  $\omega_0$  und Frequenzbreite  $\Delta\omega$ ):

$$\mathcal{F}(t) = e^{-\frac{1}{2}[(\frac{n}{c_0}x-t)\Delta\omega]^2} \sin \left[ \left( \frac{n}{c_0}x - t \right) \omega_0 \right] \quad (2.6.3)$$

<sup>42</sup>Wenn ein Feld zu schnell ansteigt, kann dies virtuelle Ladungsakkumulationen an der TF/SF-Kontur zur Folge haben, welche die Divergenzbedingung verletzt.

und dem differenzierten GAUSSschem Profil (Mittenfrequenz  $\omega_0$ ):

$$\mathcal{F}(t) = \left( \frac{n}{c_0} x - t \right) \omega_0 e^{\frac{1}{2}} e^{-\frac{1}{2}[(\frac{n}{c_0} x - t)\omega_0]^2} \quad (2.6.4)$$

Die Pulsformen  $\mathcal{F}(t)$  sind in Abbildung 2.6.4 für verschieden Parameter zusammen mit dem normierten Frequenzspektrum dargestellt.

Sollte es sich bei dem System nicht um eine Umgebung mit Grenzfläche(n) handeln, so besteht die Anregung nicht nur aus dem auf die Grenzfläche einfallendem Feld, sondern aus der Lösung des gesamten Umgebungssystems, das heißt, für den Fall der Grenzfläche, aus einfallendem, reflektiertem und transmittiertem Feld. Da sich bei Grenzflächen mit Dispersion die Pulsform ändert, und zusätzlich Modifikationen an den Hilfsdifferentialgleichungen des PMLs notwendig sind (siehe [29]), wird dieser Fall in dieser Arbeit nicht behandelt. Des Weiteren muss man bei der Berechnung der Absorptionsquerschnitte über der TF/SF-Kontur aufpassen, was diese eigentlich aussagen. Es ist daher angebracht in so einem Fall die Absorption mittels Oberflächenintegral über Partikeloberfläche(n) zu bestimmen.

#### 2.6.4. Diskrete FOURIER Transformation

Um die mittels eines Zeitschrittverfahrens gewonnenen Ergebnisse einer spektralen Analyse zu unterziehen, ist es notwendig, diese mit einer FOURIER -Transformation zu behandeln. Die in der Numerik häufig verwendete FFT<sup>43</sup> scheidet aus, da sie nur nach der Simulation angewendet werden kann. Dies ist jedoch aufgrund der dann zu speichernden Datenmengen nicht möglich. Deswegen kommt hier eine gewöhnliche diskrete FOURIER-Transformation zum Einsatz:

$$\check{F}(\omega) = \int_{t=-\inf}^{t=+\inf} f(t) e^{i\omega t} dt = \sum_{t=t_1}^{t=t_2} \Delta t f(t) e^{i\omega t} \quad (2.6.5)$$

Diese kann mittels eines Aktualisierungsalgorithmus parallel zur eigentlichen Simulation mitlaufen:

$$\check{F}(\omega)_{t+\Delta t} = \check{F}(\omega)_t + \Delta t f(t) e^{i\omega t} \quad (2.6.6)$$

Die Zeiten  $t_1$  und  $t_2$ , welche auch die Dauer der gesamten Simulation bestimmen, sind so zu wählen, das zum Zeitpunkt  $t_1$  das Feld der Anregung noch unterhalb einer Toleranz-

<sup>43</sup>eng.: „Fast-FOURIER-Transformation“ ([55])

schwelle (z.B.  $\mathcal{F} < 10^{-9}$ ) liegt, und zum Zeitpunkt  $t_2$  das Feld schon wieder hinreichend abgeklungen ist. Während sich  $t_1$  mit Hilfe der von  $\mathcal{F}$ , der Dimension des Systems und der Lichtgeschwindigkeit im Medium gut vorhersagen lässt, ist die Wahl von  $t_2$  sehr stark von der Güte des betrachteten Systems abhängig. Der Zeitverlauf des Frequenzfilters

### Zeitentwicklung des DFT-Filters

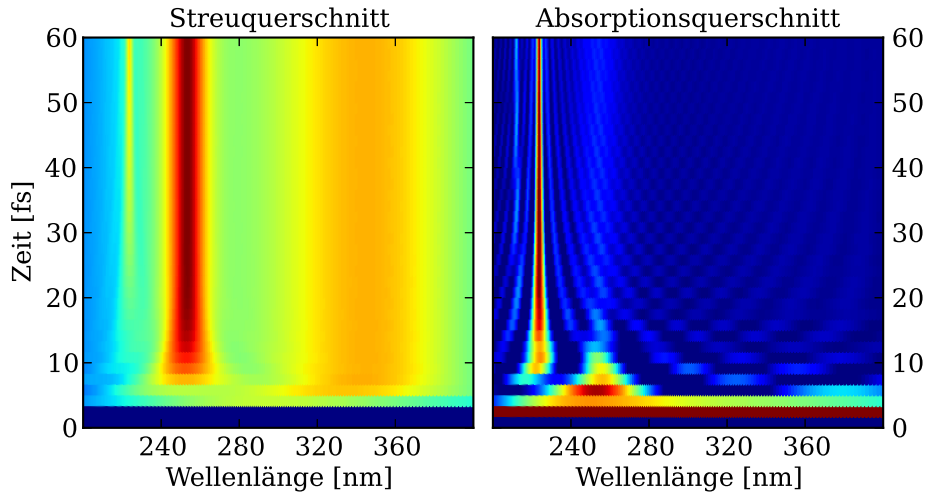


Abbildung 2.6.5.: Zeitverlauf des DFT-Filters.

### Zeitentwicklung des DFT-Filters

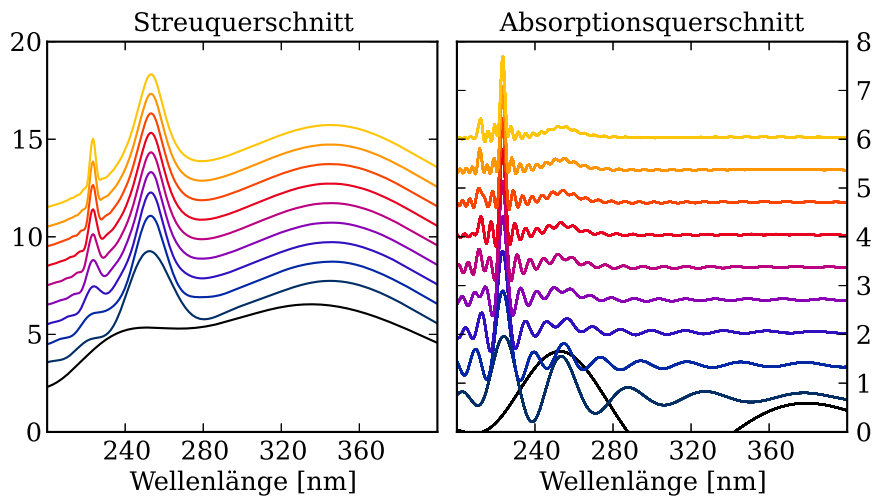


Abbildung 2.6.6.: Zeitverlauf des DFT-Filters in 5 fs Schritten für die Zeiten von 5 fs (schwarz) bis 50 fs (gelb).

$(\ddot{F}(\omega, t) = \ddot{F}(\omega)_t)$  ist für den Fall der Streu- und Absorptionquerschnitte einer 100 nm

großen Silber-Kugel (Vergleiche Kapitel 2.4.2 und Abbildung 2.4.3, 2.4.4 und 2.6.1) in den Abbildungen 2.6.5 und 2.6.6 gezeigt. Deutlich sieht man, dass der Streuquerschnitt viel schneller konvergiert, als der Absorptionsquerschnitt. Dieses Konvergenzverhalten ist ein Maß für die Zeitskalen der Streuung und Absorption im Partikel. Die Ursache liegt darin begründet, dass höhere Multipolordnungen einerseits eine höhere Güte, wie auch ein schlechteres Abstrahlverhalten zeigen, so dass die Streuprozesse im Wesentlichen innerhalb der ersten fs nach der Anregung (hier bei  $t = 1.67$  fs) abklingen.

### 2.6.5. Dimensionslose Gleichungen

In der Numerik wird in der Regel mit dimensionslosen Größen gerechnet. Dies führt dazu, dass man sich für ein Einheitssystem entscheidet, welches den numerischen Größen angepasst ist. Nun zeigt sich aber, dass die gängigen Einheitssysteme (SI-System oder GAUSS-System) nur bedingt geeignet sind, da immer wieder sehr große und sehr kleine Zahlen miteinander verrechnet werden müssen. Dies führt auf Grund der endlichen Zahlenrepräsentation im Computer (siehe [55]) zu numerischen Fehlern<sup>44</sup>. Um diese zu minimieren, wählt man ein Einheitssystem, bei dem alle relevanten Rechengrößen in der Größenordnung von 1 liegen. Die Beziehungen unter den physikalischen Größen wird dann mittels geeigneter Vorfaktoren aufrechterhalten.

Mit Hilfe der Lichtgeschwindigkeit im Vakuum  $c_0$  und der charakteristischen Länge der in dieser Arbeit verwendeten Strukturen  $l_0 = 1\text{ }\mu\text{m}$  transformieren sie die Längen und Zeiten zu:

$$\tilde{v} = \frac{1}{c_0}v \quad \tilde{r} = \frac{1}{l_0}r \quad \tilde{t} = \frac{c_0}{l_0}t \quad (2.6.7a)$$

Ebenso wird mit den Ableitungen verfahren:

$$\tilde{\partial}_t = \frac{l_0}{c_0}\partial_t \quad \tilde{\partial}_r = l_0\partial_r \quad (2.6.7b)$$

Für die elektromagnetischen Felder nutzt man die Feldstärke  $E_0 = 10^6\text{ V/m}$  und die Vakuumimpedanz  $Z_0$ :

$$\tilde{E} = \frac{1}{E_0}E \quad \tilde{H} = \frac{Z_0}{E_0}H \quad (2.6.7c)$$

Einen Sonderfall nimmt hier die Dichte  $n$  ein, welche nicht mit  $n_0 = l_0^3$  sondern mit der

---

<sup>44</sup>engl.: “Finite Precision Effects“

Ionendichte im Kristall  $n_0$  in Abhängigkeit des gewählten Materials transformiert wird:

$$\tilde{n} = \frac{1}{n_0} n \quad (2.6.7d)$$

Damit lassen sich nun die physikalischen Gleichungen in computergerechte Form bringen. Dies wird nun am Beispiel der kombinierten MAXWELL-EULER-Gleichungen für die Hydrodynamik gezeigt:

$$\frac{c_0}{l_0} n_0 \tilde{\partial}_t \tilde{n} + \frac{1}{l_0} n_0 c_0 \tilde{\nabla}(\tilde{n} \tilde{v}) = 0 \quad (2.6.8a)$$

$$\tilde{\partial}_t \tilde{n} + \tilde{\nabla}(\tilde{n} \tilde{v}) = 0 \quad (2.6.8b)$$

$$\begin{aligned} \frac{c_0}{l_0} n_0 c_0 \tilde{\partial}_t(\tilde{n} \tilde{v}) + \frac{1}{l_0} n_0 c_0 c_0 \tilde{\nabla}(\tilde{n} \tilde{v} \otimes \tilde{v}) + \frac{1}{m_e} l_0 p_0 \tilde{\nabla} \tilde{p} = \\ \frac{-e_0}{m_e} n_0 \tilde{n} (E_0 \tilde{E} + \mu c_0 \frac{E_0}{Z_0} \tilde{v} \times \tilde{H}) - \frac{c_0}{l_0} n_0 c_0 \frac{1}{\tilde{\tau}} \tilde{n} \tilde{v} \end{aligned} \quad (2.6.8c)$$

$$\begin{aligned} \tilde{\partial}_t(\tilde{n} \tilde{v}) + \tilde{\nabla}(\tilde{n} \tilde{v} \otimes \tilde{v} + \frac{l_0^2 p_0}{c_0^2 n_0 m_e} \tilde{p}) = \\ \frac{-e_0 l_0 E_0}{m_e c_0^2} \tilde{n} (\tilde{E} + \mu_r \tilde{v} \times \tilde{H}) - \frac{1}{\tilde{\tau}} \tilde{n} \tilde{v} \end{aligned} \quad (2.6.8d)$$

$$\epsilon \frac{c_0}{l_0} E_0 \tilde{\partial}_t \tilde{E} = -\frac{1}{l_0} \frac{E_0}{Z_0} \tilde{\nabla} \times \tilde{H} - (-e_0 n_0 c_0) \tilde{n} \tilde{v} \quad (2.6.8e)$$

$$\epsilon_r \tilde{\partial}_t \tilde{E} = -\tilde{\nabla} \times \tilde{H} - \frac{-e_0 l_0 n_0}{\epsilon_0 E_0} \tilde{n} \tilde{v} \quad (2.6.8f)$$

$$\mu \frac{c_0}{l_0} \frac{F_0}{Z_0} \tilde{\partial}_t \tilde{H} = \frac{1}{l_0} E_0 \tilde{\nabla} \times \tilde{E} \quad (2.6.8g)$$

$$\mu_r \tilde{\partial}_t \tilde{H} = \tilde{\nabla} \times \tilde{E} \quad (2.6.8h)$$

Dabei wurde verwendet, dass  $Z_0 = \sqrt{\mu_0/\epsilon_0}$  und  $\mu_0 \epsilon_0 = 1/c_0^2$  ist. Des Weiteren ergibt sich der Druck  $p_0$  aus  $p(n_0)$ .





### 3. Unstetiges GALERKIN-Verfahren

In diesem Kapitel geht es um das unstetige GALERKIN-Verfahren, welches in Anlehnung an die englische Bezeichnung mit DG für „Discontinuous GALERKIN“ abgekürzt wird. Dieses Verfahren wurde in den 1970ern zur Lösung partieller Differentialgleichungen (PDGL) eingeführt [56]. Es gibt mehrere mögliche Zugänge zu diesem Thema. Zum Beispiel kann man direkt von einem zugehörigen Variationsproblem ausgehen oder, wenn man gleich mit MAXWELL-Gleichungen beginnen möchte, aus diesen das DG-Verfahren ableiten. In dieser Arbeit wird die Methode anhand einer einfachen Kontinuitätsgleichung dargestellt, bevor einige Unterkapitel zur praktischen Umsetzung folgen. Schließlich wird in den folgenden Kapiteln gezeigt werden, dass sich alle hier behandelten Differentialgleichungen als eben solche Kontinuitätsgleichungen umschreiben lassen. Dabei wird im Wesentlichen wie in [15, 16] beschrieben vorgegangen.

#### 3.1. Allgemeine Herleitung

Wie erwähnt, soll hier zunächst mit einer partiellen Differentialgleichung in der Erhaltungssatzform

$$\partial_t u(\vec{r}, t) + \vec{\nabla} \cdot \vec{f}(u, \vec{r}, t) = g(\vec{r}, t) \quad \text{mit} \quad \vec{r} \in \Omega \in \mathbb{R}^3 \quad (3.1.1)$$

auf einem endlichen Gebiet  $\Omega$  begonnen werden. Dabei ist  $u(\vec{r}, t)$  die Dichte<sup>1</sup>,  $\vec{f}(u, \vec{r}, t)$  die zugehörige Flussdichte und  $g(\vec{r}, t)$  ein vorsorglich eingeführter Quellterm. Die zugehörigen Randbedingungen auf dem Rand  $\partial\Omega$  des Gebietes  $\Omega$  lauten:

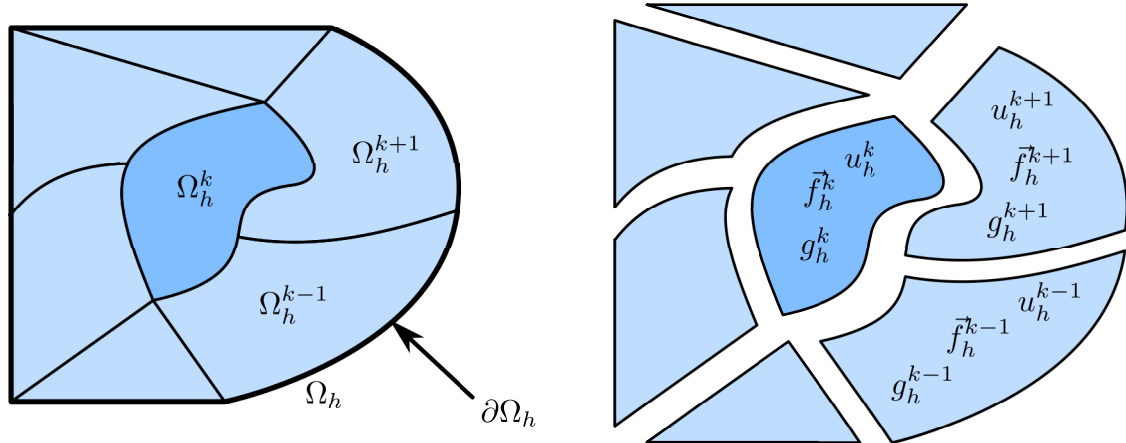
$$u(\vec{r}, t) = h_\partial(\vec{r}, t), \quad \text{mit} \quad \vec{r} \in \partial\Omega \in \mathbb{R}^3 \quad (3.1.2a)$$

$$u(\vec{r}, 0) = h_0(\vec{r}) \quad (3.1.2b)$$

wobei  $h_0$  der Startwert der PDGL ist und  $h_\partial$  die Bedingungen auf dem Rand von  $\Omega$  festlegt.

---

<sup>1</sup>Dabei ist es momentan egal, ob es sich um eine Massen-, Ladungs-, Teilchen- oder sonstige Dichte handelt.



**Abbildung 3.1.1.:** Räumliche Diskretisierung eines Gebietes  $\Omega$  in  $N_e$  Teilgebiete  $\Omega_h^k$ . Diese Gebiete können sowohl geradlinige wie auch krummlinige Ränder  $\partial\Omega_h^k$  besitzen.

Als nächstes wird das Gebiet  $\Omega$  folgendermaßen zerlegt:

$$\Omega \approx \Omega_h = \bigcup_{k=1}^{N_e} \Omega_h^k \quad (3.1.3)$$

Dabei sollen die  $N_e$  einzelnen Teilgebiete  $\Omega_h^k$  zusammenhängen, nicht überlappen und deren Vereinigung  $\Omega_h$  dem ursprünglichen Gebiet  $\Omega$  nahe kommen. Der Index  $h$  stellt ein Maß für die Feinheit der Approximation des Gebietes dar<sup>2</sup>. Ein Beispiel so einer Zerlegung ist in Abbildung 3.1.1 dargestellt. Dabei gibt es momentan keine weiteren Einschränkungen seitens der Form der Teilgebiete. Nichtsdestotrotz sollen hier auch echte Approximationen ( $\Omega \neq \Omega_h$ ) zugelassen werden, da diese im späteren Verlauf für weitere Vereinfachungen notwendig sind. Diese Teilgebiete werden im Folgenden auch Elemente genannt. Nun werden auf jedem dieser Elemente die zugehörigen Dichte-, Flussdichte- und Quellenfunktionen definiert. Dabei ergibt die direkte Summe (kanonische Fortsetzung) der Teilfunktionen die Näherung der Gesamtfunktion auf  $\Omega_h$ . Darüber hinaus werden diese Funktionen in jedem Element in eine endliche,  $N_p^k$  lange Linearkombination (LKB) zerlegt, welche aus einer ortsabhängigen Basis  $\psi_i^k(\vec{r})$  auf  $\Omega_h^k$  und zeitabhängigen Koeffizienten  $\tilde{u}_i^k(t)$ ,  $\tilde{f}_i^k(t)$  und  $\tilde{g}_i^k(t)$

<sup>2</sup>Genau beschreibt hier  $h$  die Charakteristische Größe der Elemente, wobei gilt: je kleiner  $h$ , desto kleiner die Elemente, desto mehr Elemente werden zur Approximation benötigt, das heißt:  $h \approx 1/N_e$ . Der genaue Zusammenhang zwischen  $h$  und  $N_e$  hängt vom konkreten Problem ab.

besteht. Die Anzahl der Basisfunktionen  $N_p^k$  kann von Element zu Element variieren.

$$u \approx u_h = \bigoplus_{k=1}^{N_e} u_h^k \quad u_h^k = \sum_{i=1}^{N_p^k} \tilde{u}_i^k(t) \psi_i^k(\vec{r}) \quad (3.1.4a)$$

$$\vec{f} \approx \vec{f}_h = \bigoplus_{k=1}^{N_e} \vec{f}_h^k \quad \vec{f}_h^k = \sum_{i=1}^{N_p^k} \tilde{f}_i^k(u_h^k, t) \psi_i^k(\vec{r}) \quad (3.1.4b)$$

$$g \approx g_h = \bigoplus_{k=1}^{N_e} g_h^k \quad g_h^k = \sum_{i=1}^{N_p^k} \tilde{g}_i^k(u_h^k, t) \psi_i^k(\vec{r}) \quad (3.1.4c)$$

Da nach all diesen Approximationen im Allgemeinen nicht mehr davon ausgegangen werden kann, dass Gleichung 3.1.1 exakt erfüllt wird, werden die Abweichungen zur exakten Lösung mit Hilfe des elementweisen Residuums  $R_h^k$  beschrieben:

$$\partial_t u_h^k(\vec{r}, t) + \vec{\nabla} \vec{f}_h^k(u_h^k, \vec{r}, t) - g_h^k(u_h^k, \vec{r}, t) = R_h^k \quad (3.1.5)$$

Nun gilt es, die Koeffizienten so zu bestimmen, dass das Residuum verschwindet, oder, wenn dies nicht möglich ist, es wenigstens zu minimieren. Zu diesem Zweck wird ein Satz von Testfunktionen  $\phi_h^k$  eingeführt, wobei gefordert wird, dass das Skalarprodukt dieser Testfunktionen mit dem Residuum verschwindet:

$$\int_{\Omega_h^k} R_h^k \phi_h^k = 0 \quad (3.1.6a)$$

$$\int_{\Omega_h^k} \left( \partial_t u_h^k(\vec{r}, t) + \vec{\nabla} \vec{f}_h^k(u_h^k, \vec{r}, t) - g_h^k(u_h^k, \vec{r}, t) \right) \phi_h^k(\vec{r}) = 0 \quad \forall \phi_h^k \quad (3.1.6b)$$

Fordert man nun, dass die Testfunktionen dieselbe Gestalt haben sollen wie die Lösung der PDGL<sup>3</sup>, das heißt  $\phi_h^k = \sum_{j=1}^{N_p} \tilde{\phi}_j^k \psi_j^k(\vec{r})$ , so erhält man  $N_e$  Gleichungssätze mit je  $N_p$  Gleichungen und Unbekannten<sup>4</sup>:

$$\int_{\Omega_h^k} \left( \partial_t u_h^k(\vec{r}, t) + \vec{\nabla} \vec{f}_h^k(u_h^k, \vec{r}, t) - g_h^k(u_h^k, \vec{r}, t) \right) \psi_j^k(\vec{r}) = 0 \quad \forall j, k \quad (3.1.7)$$

<sup>3</sup>Prinzipiell ist die genaue Wahl der Testfunktionen beliebig. Wählt man zum Beispiel  $N_p$  Deltadistributionen of jedem  $\Omega_h^k$  so wird die PDGL an den zugeordneten Punkten erfüllt. BORIS GALERKIN (\*1871 – †1945) führte 1915 erstmals den nach ihm benannten Ansatz ein, die Basisfunktionen auch als Testfunktionen zu verwenden. Dieser hat die Forderung zur Folge, dass das Residuum senkrecht bezüglich des gewählten Skalarproduktes auf dem Raum der Test- und Basisfunktionen steht.

<sup>4</sup>Hier wurde einfachheitshalber angenommen, dass die Anzahl der Basisfunktionen je Element gleich ist, das heißt,  $N_p^k = N_p \quad \forall k \in [1, N_e]$

An dieser Stelle ist es sinnvoll zu bemerken, dass es sich bei den  $N_e$  Gleichungssätzen um vollkommen unabhängige (ungekoppelte) Gleichungssätze handelt, das heißt, dass ein Element keine Informationen über seine Nachbarelemente kennt. Um nun die fehlenden Informationen auszutauschen, hilft die Rückbesinnung auf die Ausgangsgleichung 3.1.1. Da diese die Form einer Kontinuitätsgleichung hat, soll im Folgenden die Kopplung der Elemente mittels der Flussdichte hergestellt werden. Bildlich gesprochen: Was bei dem einen Element an Dichte abnimmt, muss in den Nachbarelementen an Dichte hinzukommen. Um diese Kopplung zu realisieren, wird zuerst partiell integriert und man erhält die sogenannte schwache Form<sup>5</sup>:

$$\begin{aligned} & \int_{\Omega_h^k} \partial_t u_h^k(\vec{r}, t) \psi_j^k(\vec{r}) - \vec{f}_h^k(u_h^k, \vec{r}, t) \vec{\nabla} \psi_j^k(\vec{r}) - g_h^k(u_h^k, \vec{r}, t) \psi_j^k(\vec{r}) \\ &= - \int_{\partial\Omega_h^k} \vec{n} \cdot \vec{f}^* \psi_j^k(\vec{r}) \quad \forall j \end{aligned} \quad (3.1.8)$$

Dabei ist  $\vec{n}$  der Flächennormalenvektor auf die betrachtete Elementgrenze. Des Weiteren wurde die echte Flussdichte  $\vec{f}_h^k$  durch eine numerische Flussdichte  $\vec{f}^*$  ersetzt. Mit Hilfe dieser numerischen Flussdichte wird nun eine Kopplung der Elemente herbeigeführt. Dabei macht man es sich zu nutze, dass an jeder Elementgrenze zwei Flussdichten existieren, einerseits die Flussdichte des gerade betrachteten Elementes und andererseits die Flussdichte des der betrachteten Fläche gegenüberliegenden Elementes. Neben der schwachen Form der PDGL gibt es auch noch die starke Form, welche jetzt noch schnell nachgereicht werden soll, bevor im Folgenden der Term mit der numerischen Flussdichte sowie selbige genauer betrachtet werden. Diese starke Form erhält man durch nochmalige partielle Integration, nun jedoch ohne die Flussdichte  $\vec{f}_h^k$  zu ersetzen:

$$\begin{aligned} & \int_{\Omega_h^k} \left( \partial_t u_h^k(\vec{r}, t) + \vec{\nabla} \vec{f}_h^k(u_h^k, \vec{r}, t) - g_h^k(u_h^k, \vec{r}, t) \right) \psi_j^k(\vec{r}) \\ &= \int_{\partial\Omega_h^k} \vec{n} \cdot \left( \vec{f}_h^k(u_h^k, \vec{r}, t) - \vec{f}^*(u_h^k, \vec{r}, t) \right) \psi_j^k(\vec{r}) \quad \forall j \end{aligned} \quad (3.1.9)$$

Die numerische Flussdichte ist für sich genommen ein eigenes Untersuchungsgebiet, da sie, wie bereits erwähnt, die Kopplung zwischen den Elementen sicherstellen soll. Des Weiteren fordert man, dass sie möglichst wenig zusätzliche Fehler in das Gleichungssystem einbringen soll, d. h., die Flussdichte muss so gewählt werden, dass die Lösung sowohl

---

<sup>5</sup>Bei der schwachen Form ist der Ortsdifferentialoperator von den Basisfunktionen auf die Testfunktionen übergegangen. Man bräuchte nun keine differenzierbaren Lösungsfunktionen mehr, was hier, da Basis- und Testfunktionen gleich sind, aber keine größere Rolle spielt.

konvergent<sup>6</sup> als auch konsistent<sup>7</sup> ist. So existieren für verschieden Anwendungsfälle verschiedene numerische Flussdichten mit den zugehörigen Vor- und Nachteilen [16]. Für das weitere Verständnis wird hier der einfachste Fall der numerischen Flussdichte, die Mittelwertsflussdichte (engl.: Central Flux), angenommen, welcher aus dem Mittelwert der links- und rechtsseitigen Flussdichte besteht<sup>8</sup>:

$$\vec{f}^* = \frac{1}{2} \left( \vec{f}_h^k + \vec{f}_h^{k+\Delta k} \right) \quad (3.1.10)$$

Hier ist  $\vec{f}_h^{k+\Delta k}$  die zur jeweiligen Fläche gehörige Flussdichte des Nachbarelementes. Damit lässt sich nun die Wirkung des zusätzlichen Terms weiter herausarbeiten. Zuerst fällt auf, dass in Gleichung 3.1.10 weder räumliche noch zeitliche Ableitungen enthalten sind. Damit kann man die Mittelwertsflussdichte als zusätzlichen Quellterm interpretieren. Wenn nun die beiden beteiligten Flussdichten, aus denen die numerische Flussdichte zusammengesetzt ist, gleich sind<sup>9</sup>, das heißt, der Mittelwert  $\vec{f}^*$  entspricht dann zwangsläufig  $\vec{f}_h^k$ , verschwindet der gesamte Integrand  $\vec{n} \cdot (\vec{f}_h^k - \vec{f}^*) = 0$  und somit auch der Kopplungsterm. Des Weiteren ist dieser Term umso stärker, je weiter die benachbarten Flussdichten auseinander liegen.

Für die folgende numerische Behandlung der Gleichungen bietet es sich an, diese in Matrixform zu bringen. Dabei werden die Integrale als Einträge von Matrizen aufgefasst<sup>10</sup>.

$$M_{ij}^k \partial_t \tilde{u}_i^k(t) + \vec{S}_{ij}^k \tilde{f}_i^k(u_h^k, t) = M_{ij}^k \tilde{g}_i^k + F_{ij}^k \left[ \vec{n} \cdot \left( \tilde{f}_i^k(u_h^k, t) - \tilde{f}_i^*(u_h, t) \right) \right] \quad (3.1.11a)$$

$$M_{ij}^k = \int_{\Omega_h^k} \psi_i^k(\vec{r}) \psi_j^k(\vec{r}) \quad (3.1.11b)$$

$$\vec{S}_{ij}^k = \int_{\Omega_h^k} \vec{\nabla} \psi_i^k(\vec{r}) \psi_j^k(\vec{r}) \quad (3.1.11c)$$

$$F_{ij}^k = \int_{\partial\Omega_h^k} \psi_i^k(\vec{r}) \psi_j^k(\vec{r}) \quad (3.1.11d)$$

Hier bezeichnet nun  $M_{ij}^k$  die Massenmatrix,  $\vec{S}_{ij}^k$  die Steifheitsmatrix<sup>11</sup> und  $F_{ij}^k$  die Ober-

<sup>6</sup>Mit feiner werdender Diskretisierung und/oder mehr Basisfunktionen wird der Fehler kleiner.

<sup>7</sup>Wenn man die exakte Lösung einsetzt, soll die numerische Flussdichte der analytischen entsprechen.

<sup>8</sup>Wer einen etwas mathematischeren Zugang wünscht, dem kann z.B. [16] nahegelegt werden.

<sup>9</sup>Gleich meint hier Betrag und Richtung.

<sup>10</sup>Hier wird wieder vereinfachend angenommen, dass  $N_p^k = N_p$  ist, d. h. die Anzahl der Basisfunktionen für jedes Element gleich ist. Diese Forderung kann man auch weglassen, jedoch muss dann im Oberflächenintegralterm mit zusätzlichen Projektoren gearbeitet werden. Näheres dazu findet sich im Kapitel 4 über die krummlinigen Elemente.

<sup>11</sup>Hier liest sich:  $\vec{S}^k \tilde{f}^k = S_x^k \tilde{f}_x^k + S_y^k \tilde{f}_y^k + S_z^k \tilde{f}_z^k$  mit  $S_{ij,x}^k = \int \partial_x \psi_i \psi_j$  und den anderen Koordinaten entsprechend.

flächenmassenmatrix<sup>12</sup>. Genauso wie für die starke Form existiert auch für die schwache Form eine derartige Matrixdarstellung, diese lautet mit obigen Bezeichnungen:

$$M_{ij}^k \partial_t \tilde{u}_i^k(t) - (S_{ij}^k)^T \tilde{f}_i^k(u_h^k, t) = M_{ij}^k \tilde{g}_i^k - F_{ij}^k \left( \vec{n} \cdot \tilde{f}_i^*(u_h^k, t) \right) \quad (3.1.12)$$

Speziell die Matrixdarstellung der starken Form lässt sich zusätzlich etwas übersichtlicher gestalten, wenn man noch die schwache Ableitungsmatrix  $\vec{D}_w^k$  einführt:

$$\vec{D}_w^k = (M^k)^{-1} \left( \vec{S}^k \right)^T \quad (3.1.13a)$$

$$\partial_t \tilde{u}^k(t) = \vec{D}_w^k \tilde{f}^k(u_h^k, t) + \tilde{g}^k - (M^k)^{-1} F^k \left[ \vec{n} \cdot \left( \tilde{f}^k(u_h^k, t) - \tilde{f}^*(u_h^k, t) \right) \right] \quad (3.1.13b)$$

Schlussendlich erhält man die Kontinuitätsgleichung in diskreter Form, bestehend aus der Zeitableitung, der Divergenz und dem Quellterm (welcher sich selbst in den physikalischen und den numerischen Teil aufspaltet). Zu beachten ist, dass die Massenmatrix, welche es zu invertieren gilt, lokal beziehungsweise als globales Gleichungssystem betrachtet, streng blockdiagonal ist und sich so sehr leicht invertieren lässt (z.B. elementweise)<sup>13</sup>.

<sup>12</sup>Diese Namen stammen aus der Anwendung der kontinuierlichen GALERKIN-Methode (klassische Finite-Elemente-Methode) auf die Kontinuumsmechanik.

<sup>13</sup>Diese Eigenschaft geht verloren, wenn man statt der Erhaltung der Dichte die Stetigkeit fordern würde. Dann würden Basisfunktionen über die Elementgrenzen hinweg, also in mehr als einem Element, existieren, so dass die globale Massenmatrix nicht mehr elementweise invertiert werden kann ([21]).

### 3.1.1. Basisfunktionen

Nachdem im vorangegangenen Kapitel die Grundzüge der DG-Methode skizziert wurden, wird es in diesem Abschnitt vorrangig um die Konstruktion eines geeigneten Basisfunktionssystems gehen. Dabei werden recht schnell einige Einschränkungen bezüglich der Elementform (der  $\Omega_h^k$ ) notwendig, die dazu führen, dass die Zerlegung des Simulationsgebietes  $\Omega$  eine echte Approximation  $\Omega \approx \Omega_h \neq \Omega$  nach sich zieht.

Mit den obigen Variablen  $N_e$  für die Elementanzahl und  $N_p$  für die Anzahl der Basisfunktionen je Element sind insgesamt  $N_e \cdot N_p^2$  verschiedene Integrale für jede der fünf Matrizen  $M$ ,  $D_x$ ,  $D_y$ ,  $D_z$  und  $F$  zu berechnen<sup>14</sup>. Um dieser recht großen Zahl Herr zu werden, ist es üblich, einerseits alle Berechnungen nur für ein Referenzelement  $I$  durchzuführen und die so gewonnenen Ergebnisse auf alle physikalischen Elemente  $\Omega_h^k$  zu transformieren, andererseits einen Basisfunktionssatz so zu wählen, dass diese Berechnungen besonders einfach durchgeführt werden können. Im Folgenden soll daher zuerst das Referenzelement und anschließend ein geeigneter Basisfunktionssatz festgelegt werden.

#### Das Referenzelement

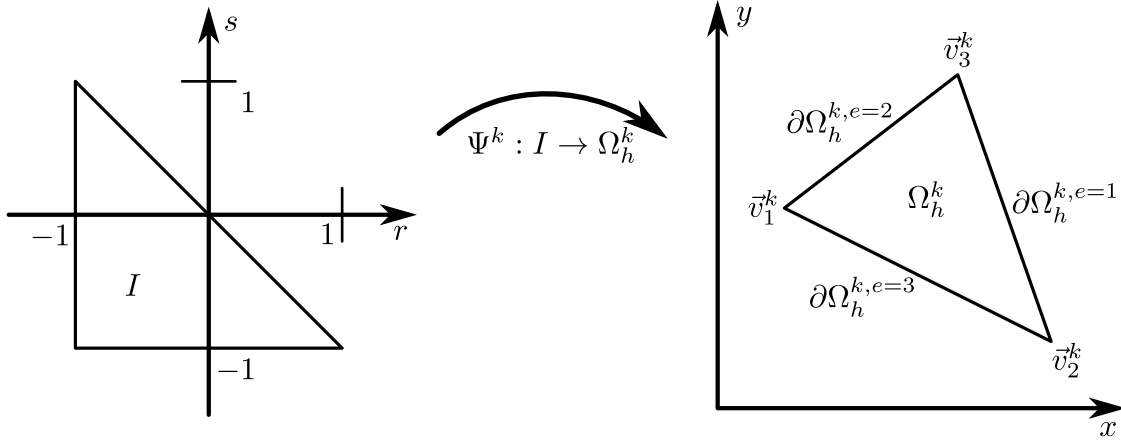
Zum weiteren Verständnis ist es hilfreich, sich mit einigen Konventionen vertraut zu machen. So bezeichnen von nun an  $r$ ,  $s$  und  $t$  die Koordinaten im Referenzelement  $I$  und  $x$ ,  $y$  und  $z$  die Koordinaten im physikalischen Element  $\Omega_h^k$ . Damit die Implementation der Koordinatentransformation  $\Psi^k : I \rightarrow \Omega_h^k$  sowie der Matrixoperatoren besonders leicht gelingt, soll die JACOBI-Matrix  $J^k$  für jedes Element  $\Omega_h^k$  konstant sein (siehe hierzu Abschnitt 3.1.2), das heißt:

$$J^k = \frac{\partial(x, y, z)}{\partial(r, s, t)} \stackrel{!}{=} \text{konst.} \quad \forall (x, y, z) \in \Omega_h^k \quad \forall (r, s, t) \in I \quad (3.1.14)$$

Diese Eigenschaft besitzen unter anderem alle sogenannten  $d$ -Simplexe (Linie, Dreieck, Tetraeder) jeweils untereinander ([15], [16]). Damit lassen sich alle  $d$ -Simplexe gleicher Dimension mit einer einfachen linearen Abbildung verknüpfen [15]<sup>15</sup>. Dies erklärt nun auch die Beliebtheit von Dreiecken und Tetraedern als Grundlage der Ortsdiskretisierung. Nichtsdestotrotz wären auch andere Referenzelemente, wie z.B. im 2D Vierecke und im 3D Pyramiden, Prismen und Hexaeder, möglich. Bei diesen ist jedoch die JACOBI-Matrix im Allgemeinen nicht mehr konstant, d. h., nicht nur vom Element, sondern auch noch vom

<sup>14</sup>Nimmt man z.B. ein Projekt mit 20000 Elementen und 35 Basisfunktionen je Element, ergeben sich daraus 122.5 Millionen Integrale.

<sup>15</sup>Durch diese Vereinfachung müssen nur die Operatoren für das Referenzelement berechnet werden. Damit ergeben sich bei obigen Beispiel noch  $N_p^2 = 1225$  Integrale.



**Abbildung 3.1.2.:** Abbildung  $\Psi^k$  vom Referenzelement  $I$  zum physikalisches Element  $\Omega_h^k$ .  $\partial\Omega_h^{k,e}$  bezeichnet die Ränder von  $\Omega_h^k$ .

Ort abhängig. Wie sich dies speziell bei krummlinig begrenzten Körpern auswirkt, wird in Kapitel 4 behandelt. Wie man sich die Transformation vorstellen könnte, ist in Abbildung 3.1.2 anhand von Dreiecken im 2D dargestellt.

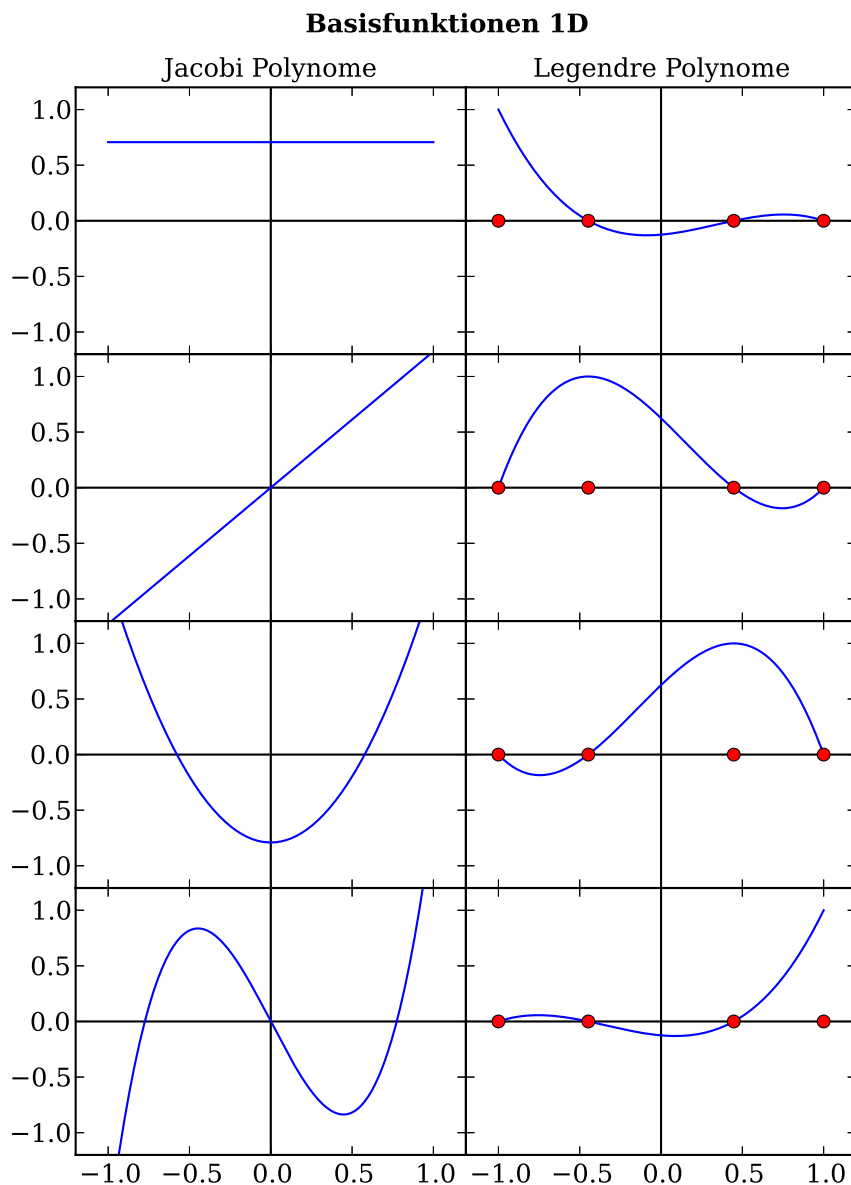
### Die Basisfunktionen

Nachdem im vorangegangenen Abschnitt das Referenzelement festgelegt wurde (je nach Dimension: Linie, Dreieck, Tetraeder), ist das Ziel dieses Abschnittes, ein geeignetes Basisfunktionssystem ausfindig zu machen. Hierzu sei speziell auf die Referenzen [15, 16, 67] verwiesen, die diesem Abschnitt zugrunde liegen. Wie bereits bei vorangegangenen Entscheidungen hat man auch hier wieder verschiedene Wahlmöglichkeiten. Eine davon ist ein orthonormiertes Basisfunktionssystem zu wählen, bei dem bekanntlich gilt:

$$M_{ij}^I = \int_I \psi_i^I \psi_j^I = \delta_{ij} \quad (3.1.15)$$

Damit entfällt zumindest die Integration der Massenmatrix ( $M^I = \mathbb{I}$ ) sowie deren Inversion, da ( $\mathbb{I}^{-1} = \mathbb{I}$ ) ist. Ein System, was diese Eigenschaften besitzt, sind die orthonormierten JACOBI-Polynome  $\psi_i$ . Einen Nachteil hat dieses System. So sind zur Auswertung des Oberflächenintegrals  $\int_{\partial_I}$  alle Basisfunktionen notwendig, obgleich der Hauptteil der Funktionen im Inneren von  $I$  liegt. Es wäre doch denkbar, zu einem System überzugehen, welches eine Aufteilung in „innere“ und „äußere“ Basisfunktionen vornimmt, so dass zur Auswertung des Integrals nur noch die „äußeren“ Basisfunktionen berücksichtigt werden müssen. Eines dieser Systeme ist der Übergang von einer modalen Basis zu einer nodalen Basis, wie sie z.B. die LEGENDRESchen Interpolationspolynome  $l_i$  mit den Interpolationsknoten  $\vec{r}_i$  dar-





**Abbildung 3.1.3.:** Vergleich der Basisfunktionen im 1D. Linke Seite: ortho-normierte JACOBI-Polynome, rechte Seite: LEGENDRE-Interpolationspolynome, welche an genau einem Interpolationspunkt Eins, und an allen anderen Null annehmen. Man erkennt die Symmetrie der Interpolationspunkte sowie deren nicht äquidistante Verteilung.

stellen. Bei dieser Basis, welche für den 1D-Fall in Abbildung 3.1.3 dargestellt ist, sind die Polynome so konstruiert, dass sie am zugehörigen Knoten Eins sind, und an den anderen Null ergeben<sup>16</sup>:

$$l_i^I(\vec{r}_j) = \delta_{ij} \quad \forall \vec{r}_j \quad (3.1.16)$$

Mit diesem Funktionssystem müssen für das Oberflächenintegral nur noch die zu den Randknoten gehörenden Basisfunktionen (beziehungsweise deren Koeffizienten) ausgewertet werden, da die inneren Basisfunktionen, das heißt, die übrigen, nicht zum Randknoten gehörenden Basisfunktionen, keinen Beitrag liefern. Eine weitere Besonderheit ist, dass die Koeffizienten der den Knoten zugeordneten Funktionen identisch mit dem Funktionswert an diesen Knoten sind, das heißt, die Polynome interpolieren eine durch die Werte an den Knoten vorgegebene Funktion. Der Zusammenhang zwischen Ordnung  $N$  der Polynome und Anzahl der Basisfunktionen  $N_p$  je Element (und damit Anzahl an Knoten je Element) errechnet sich zu:

$$N_p^{1D} = N + 1 \quad (3.1.17a)$$

$$N_p^{2D} = \frac{(N+1)(N+2)}{2} \quad (3.1.17b)$$

$$N_p^{3D} = \frac{(N+1)(N+2)(N+3)}{6} \quad (3.1.17c)$$

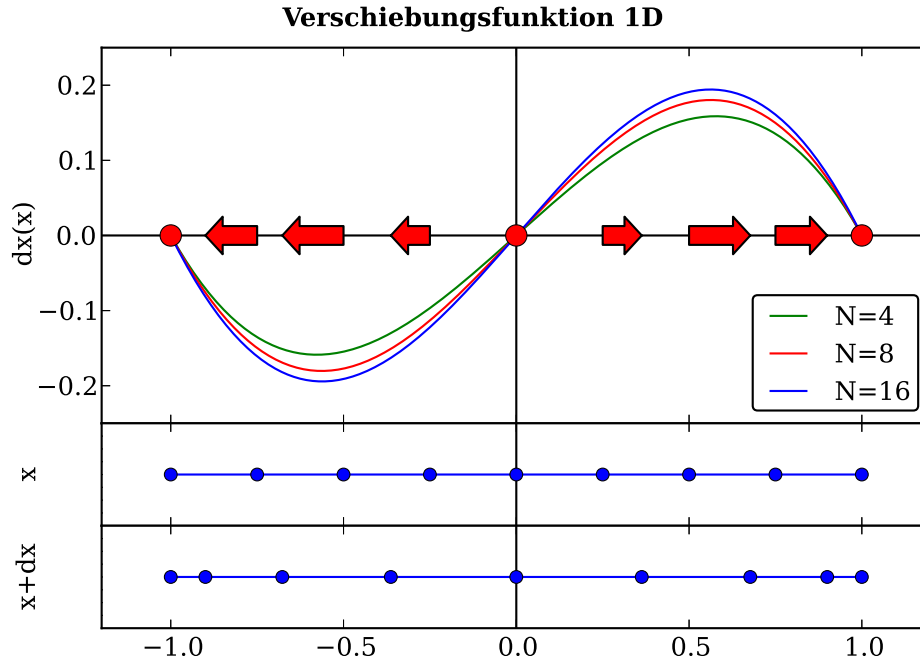
und ist in Tabelle 3.1.1 für die ersten acht Ordnungen aufgeführt. Die Anzahl  $N_{fp}$  der Knoten je Fläche  $\partial\Omega_h^{k,e}$  entspricht der Knotenanzahl der jeweils niedrigeren Dimension. Es bleibt noch die Frage nach der Position der passenden Interpolationsknoten zu be-

**Tabelle 3.1.1.:** Anzahl der Knoten  $N_p$  im d-Simplex in Abhängigkeit der Ordnung  $N$  und der Dimension.

N	1	2	3	4	5	6	7	8
1D	2	3	4	5	6	7	8	9
2D	3	6	10	15	21	28	36	45
3D	4	10	20	35	56	84	120	165

antworten, da diese die LEGENDRE-Polynome festlegen. Dazu wird eine Abbildung von den orthonormierten JACOBI-Polynomen zu den LEGENDRE-Polynomen definiert mit dem Ziel, die vorteilhaften Eigenschaften beider Systeme, das heißt die, einfache Volumenintegration der JACOBI-Polynome und die Auswertung des Oberflächenintegrals mittels der

<sup>16</sup>Obgleich die Funktionen am zugehörigen Knoten Eins sind, heißt das weder, dass das Maximum Eins sein, noch dass das Maximum an einem der Interpolationsknoten liegen muss.



**Abbildung 3.1.4.:** Verschiebungsfunktion  $w(x)^{1D}$  im 1D, als Differenz zwischen äquidistanter Knotenverteilung und den LGL-Punkten. Erste Zeile: Resultierende Verschiebungsfunktion für verschiedene Ordnungen sowie durch Pfeile skizzierte Verschiebung am Beispiel der 8. Ordnung. Zweite Zeile: Äquidistante Knotenverteilung 8. Ordnung. Dritte Zeile: LGL-Punkte 8. Ordnung.

LEGENDRE-Polynome, miteinander kombinieren zu können:

$$\sum \tilde{u}_i^I \psi_i^I(\vec{r}) = \sum u_i^I l_i^I(\vec{r}) \quad (3.1.18a)$$

$$u_i^I = u_h^I(\vec{r}_i) \quad (3.1.18b)$$

Damit ergibt sich die Transformationsmatrix  $V^I : \tilde{u}_i^I \rightarrow u_i^I$ , in Zukunft VANDERMONDE-Matrix genannt, zu:

$$V^I \tilde{u}^I = u^I \quad (3.1.19a)$$

$$V_{ij}^I = \psi_j^I(\vec{r}_i) \quad (3.1.19b)$$

Ziel ist nun, diese Transformationsmatrix so numerisch stabil wie möglich zu gestalten. Dabei ist nach Gleichung 3.1.19 mit der Festlegung auf die JACOBI-Polynome  $\psi_j$  die Verteilung der Knotenpunkte  $\vec{r}_i$  der einzige noch verbliebene Freiheitsgrad. Hierbei stellt sich heraus, dass eine äquidistante Verteilung eine suboptimale Wahl darstellt. Eine geeignete Verteilung kann man für 1D-Elemente mit den LEGENDRE-GAUSS-LOBATTO Punkten

(LGL) als Interpolationsknoten erreichen<sup>17</sup>. Leider sind derartige Regeln für den höherdimensionalen Raum bislang nicht bekannt, so dass hier der Ansatz aus [67] zum Tragen kommt, der hier kurz skizziert werden soll. Dazu definiert man sich eine Verschiebungs-/Projektionsfunktion  $w(x)^{1D}$ , welche die „schlechten“ Punkte in „gute“ Punkte verschiebt. Dies ist in Abbildung 3.1.4 anhand der 8. Ordnung skizziert, und für zwei weitere Ordnungen dargestellt. Dabei fällt auf, dass sich die Verschiebungsfunktion mit der Ordnung nicht sehr stark ändert. Für den 2D-Fall wird diese, ordnungsabhängige, Verschiebungsfunktion an den Kanten des Dreiecks angelegt und in die Fläche hineingezogen. Dafür geht man von den „gewöhnlichen“ Koordinaten  $(r, s)$  auf die baryzentrischen Koordinaten  $(\lambda_1, \lambda_2, \lambda_3)$  über. Diese haben folgende Eigenschaften:

$$0 \leq \lambda_i \leq 1, \quad \sum_i \lambda_i = 1 \quad (3.1.20)$$

Damit lässt sich jeder Punkt  $\vec{r}$  im Dreieck mithilfe dieser Koordinaten und den Eckvertices  $\vec{v}_i$  beschreiben:

$$\vec{r} = \sum_i \vec{v}_i \lambda_i \quad (3.1.21)$$

Das heißt, dass jede baryzentrische Koordinate Eins am zugehörigen Eckvertex und Null an der gegenüberliegenden Kante ist. Die baryzentrischen Koordinaten lassen sich als Funktion von  $(r, s)$  angeben:

$$\lambda_1(r, s) = \frac{1+s}{2} \quad (3.1.22a)$$

$$\lambda_2(r, s) = -\frac{r+s}{2} \quad (3.1.22b)$$

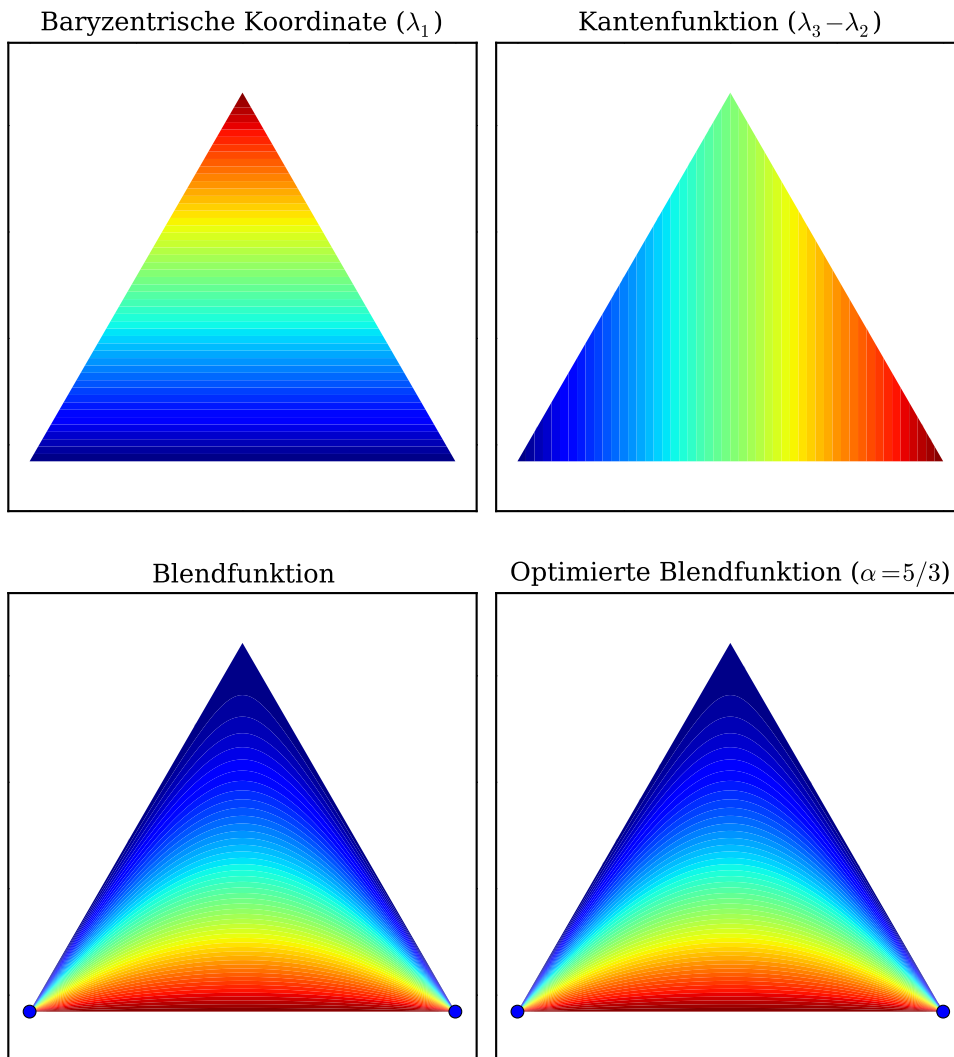
$$\lambda_3(r, s) = \frac{1+r}{2} \quad (3.1.22c)$$

Mit Hilfe dieser Koordinaten lässt sich nun eine Kantenfunktion  $x^{2D}(r, s)$  konstruieren die entlang einer Kante des Dreiecks läuft, und gleichzeitig im ganzen Dreieck definiert ist. In der ersten Zeile von Abbildung 3.1.5 ist diese Kantenfunktion für  $(e = 1)$  zusammen mit der baryzentrischen Koordinate  $\lambda_1$  dargestellt. Damit wird aus der eindimensionalen Verschiebungsfunktion  $w(x)$  die Kantenverschiebungsfunktion (für die Kante  $e = 1$ ):

$$\vec{w}_1^{2D}(r, s) = w^{1D}(\lambda_3 - \lambda_2) \vec{t}_1 \quad (3.1.23)$$

Hier wurde zusätzlich der Vektor entlang der Kante  $(e = 1)$   $\vec{t}_1 = (\vec{v}_3 - \vec{v}_2)/|\vec{v}_3 - \vec{v}_2|$

<sup>17</sup>Die LEGENDRE-GAUSS-LOBATTO-Punkte sind die GAUSS-Quadraturpunkte erweitert um die Randpunkte des betrachteten Intervalls.

**Baryzentrische Koordinaten & Blendfunktion 2D**

**Abbildung 3.1.5.:** Darstellung der baryzentrischen Koordinate für  $\lambda_1$ , der „Kantenfunktion“  $\lambda_3 - \lambda_2$ , der Blendfunktion und der optimierten Blendfunktion. An den blauen Punkten ist die Blendfunktion nicht definiert, da sie dort Null und Eins gleichzeitig sein müsste. Da die Verschiebungsfunktion an diesen Stellen verschwindet, kann dieses Verhalten vernachlässigt werden.

genutzt, um aus der skalaren Verschiebungsfunktion eine Vektorverschiebungsfunktion zu gewinnen. Die zugehörige „Blendfunktion“, welche  $\vec{w}_1^{2D}$  ausblendet, je weiter ein Punkt von der betreffenden Kante entfernt ist, lautet:

$$b_1^{2D}(r, s) = \frac{4\lambda_3\lambda_2}{1 - (\lambda_3 - \lambda_2)^2} \quad (3.1.24)$$

Für eine weitere Optimierung des Knotensatzes kann man zu einer modifizierten „Blendfunktion“ übergehen, welche einen Variationsparameter  $c_\alpha$  enthält

$$\tilde{b}_1^{2D}(r, s) = (1 + (c_\alpha\lambda_1)^2) \frac{4\lambda_3\lambda_2}{1 - (\lambda_3 - \lambda_2)^2} \quad (3.1.25)$$

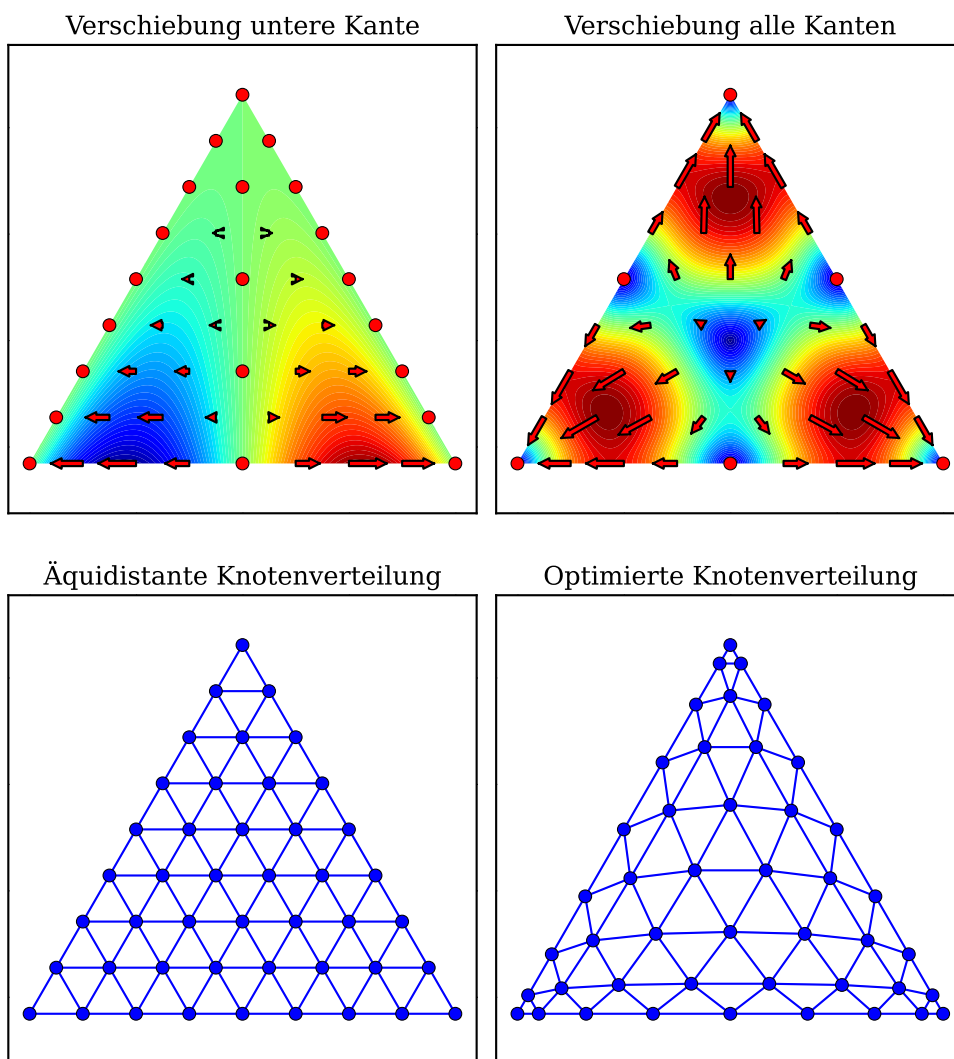
Diese Blendfunktionen sind in der zweiten Zeile von Abbildung 3.1.5 gegenübergestellt. Die optimierte Vektorverschiebung ( $\vec{g}_1^{2D}$ ) entlang der Kante (e=1) ergibt sich nun zu:

$$\vec{g}_1^{2D} = \tilde{b}_1^{2D} \vec{w}_1^{2D} \quad (3.1.26)$$

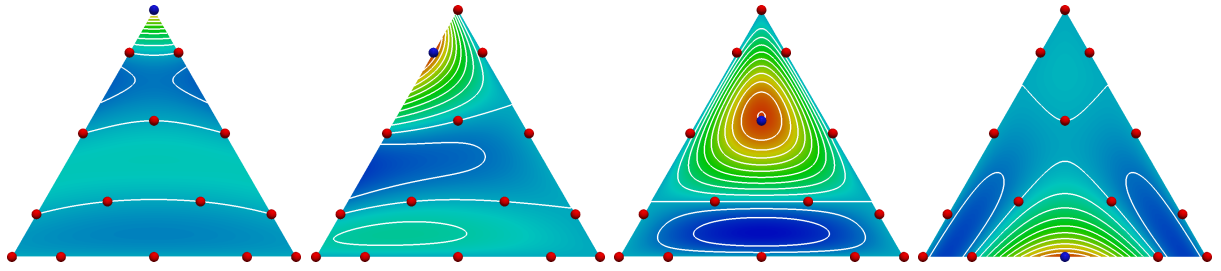
Und die Gesamtverschiebungsfunktion  $\vec{g}^{2D}$  als Summe der Einzelverschiebungsfunktionen:

$$\vec{g}^{2D} = \sum_i \vec{g}_i^{2D} \quad (3.1.27)$$

Dies ist in Abbildung 3.1.6 sowohl für eine Kante, wie auch als Summe über alle Kanten dargestellt. Mit Hilfe dieser so konstruierten Verschiebung werden im Anschluss wieder die äquidistanten Knoten in optimierte Knoten überführt (siehe Abbildung 3.1.6 für die 8. Ordnung). Im Dreidimensionalen wird analog vorgegangen, nur dass nun zwei Verschiebungsschritte vorgenommen werden müssen: Der Erste von der Kante zur Fläche und der Zweite von der Fläche zum Volumen. Mithilfe der so gewonnenen Knotenpunkte, den ortho-normierten Polynomen im Referenzelement und der Definition der VANDERMONDEmatrix (siehe Gleichung 3.1.19) lassen sich nun die Interpolationspolynome konstruieren. Diese sind für ein 2D-Element mit einer Approximationsordnung  $N = 4$  in Abbildung 3.1.7 dargestellt.

**Verschiebungsfunktion 2D**

**Abbildung 3.1.6.:** Konstruktion der Volumen-Verschiebungsfunktion, ausgehend von einer Kante, sowie die Verteilung der Knotenpunkte vor und nach der Verschiebung. Die Farben zeigen den Betrag der Verschiebungsfunktion.



**Abbildung 3.1.7.:** Basisfunktionen im Zweidimensionalen auf einem  $\Omega_h^k$ , definiert durch Interpolationspolynome, welche 1 an dem blauen Punkt und 0 an den roten Punkten sind. Man erkennt die nicht äquidistante Verteilung der Knotenpunkte. Des Weiteren ist das Maximum der 3. Basisfunktion größer Eins. Die Konturlinien haben einen Abstand von 0.1 zueinander. Dargestellt sind nur 4 der 15 Basisfunktionen, da sich die verbleibenden durch Drehung und Spiegelung der dargestellten Funktionen ergeben.

### 3.1.2. Lokale Operatoren

Nachdem die Wahl der Basis mit den Interpolationspolynomen unter Zuhilfenahme der orthonormierten Polynome im Referenzelement sowie der Referenzelemente an sich abgeschlossen ist, soll es nun um die lokalen Operatoren  $M$ ,  $\vec{D}$  und  $F$  gehen. Dabei wird zunächst die Massenmatrix betrachtet:

$$M_{ij}^k = \int_{\Omega_h^k} l_i^k l_j^k = \int_I J^k l_i^I l_j^I \quad (3.1.28)$$

Hier bezeichnet zur Vereinfachung der Schreibweise  $J^k$  gleich die JACOBI-Determinante. Des Weiteren wird jetzt ersichtlich, wieso bei der Wahl der Elemente eine möglichst einfache Transformation im Vordergrund stand, denn nun kann man  $J^k$ , da es sich nur noch um eine Zahl handelt, aus dem Integral herausziehen. Wendet man noch obige VANDERMONDE-Matrix geschickt an, erhält man:

$$M^k = J^k (VV^T)^{-1} \quad (3.1.29)$$

Um die nun hinzugekommene Inversion muss man sich nicht weiter kümmern, da die Massenmatrix in den Gleichungen ihrerseits nur invertiert vorkommt, so dass schließlich gar keine Inversion mehr nötig ist. Etwas aufwendiger ist die Berechnung der Ableitungsmatrix  $\vec{D} = M^{-1}\vec{S}$ . Dazu wird eine Hilfsmatrix  $\vec{V}_I$  eingeführt:

$$\vec{V}_I = \vec{D}_I V \quad (3.1.30)$$



Der Index  $I$  verweist darauf, dass hier die Ableitung nach den Koordinaten des Referenzelementes  $I$  gebildet wird. Diese enthält folgende Einträge:

$$\vec{V}_{I,ij} = \vec{\nabla}_I \psi_j(\vec{r}_i) \quad (3.1.31)$$

welche sich leicht über entsprechende Rekursionsformeln ermitteln lassen, so dass sich die Ableitungsmatrix zu

$$\vec{D}_I = \vec{V}_I V^{-1} \quad (3.1.32)$$

ergibt. Es ist an dieser Stelle wichtig darauf hinzuweisen, dass  $\vec{D}_I$  hier nicht mehr vom konkreten Element  $\Omega_h^k$  abhängt. Die Ableitungsmatrix  $\vec{D}^k$  nach den physikalischen Koordinaten  $x$ ,  $y$  und  $z$  erhält man schließlich mit:

$$D_x^k = \left( \frac{\partial r}{\partial x} \right)^k D_r + \left( \frac{\partial s}{\partial x} \right)^k D_s + \left( \frac{\partial t}{\partial x} \right)^k D_t \quad (3.1.33a)$$

$$D_y^k = \left( \frac{\partial r}{\partial y} \right)^k D_r + \left( \frac{\partial s}{\partial y} \right)^k D_s + \left( \frac{\partial t}{\partial y} \right)^k D_t \quad (3.1.33b)$$

$$D_z^k = \left( \frac{\partial r}{\partial z} \right)^k D_r + \left( \frac{\partial s}{\partial z} \right)^k D_s + \left( \frac{\partial t}{\partial z} \right)^k D_t \quad (3.1.33c)$$

Hier sind die partiellen Ableitungen nach den Koordinaten selbstverständlich vom konkreten physikalischen Element  $\Omega_h^k$  abhängig. Diese Quotienten sind im Allgemeinen, und speziell für gekrümmte Elemente, auch vom Ort abhängig, so dass sie nur für spezielle Elementtypen (z.B. Dreiecke und Tetraeder) für jedes Element konstant sind und sich so als einfache Zahl ausdrücken lassen. Neben der „starken“ Ableitung wird für die schwache Form der PDGL noch die „schwache“ Ableitung  $\vec{D}^w$  benötigt.

$$\vec{D}_I^w = M^{-1} \vec{S}_I^T \quad (3.1.34)$$

$$= V V^T (V^T)^{-1} \vec{V}_I^T M^T \quad (3.1.35)$$

$$= V \vec{V}_I^T M \quad (3.1.36)$$

Bei der letzten Umformung ist zu beachten, dass die Massenmatrix  $M$  symmetrisch ist, das heißt,  $M = M^T$ . Bleibt als letzter Operator nur noch die Oberflächenmassenmatrix  $F$ . Einfachheitshalber wird im Folgenden nur eine Nachbarfläche betrachtet.  $F$  wird analog zur Volumenmassenmatrix  $M^k$  bestimmt, mit dem Unterschied, dass nun neben dem

Elementindex  $k$  noch der Flächenindex  $e$  hinzukommt (siehe Abbildung 3.1.2).

$$M_{ij}^{k,e} = \int_{\partial\Omega_h^{k,e}} l_j^k l_i^k \quad (3.1.37a)$$

$$M^{k,e} = J^{k,e} \left( V^{k,e} (V^{k,e})^T \right)^{-1} \quad (3.1.37b)$$

Die gesamte Flächenmassenmatrix ergibt sich dann als Summe der einzelnen Beiträge (mit  $N_f$  als Anzahl der Oberflächen):

$$F^k = \sum_{e=1}^{N_f} M^{k,e} \quad (3.1.38)$$

Neben diesen Hauptoperatoren ist es noch hilfreich, zwei spezielle Matrizen, die Filtermatrix  $FI$  und die Interpolationsmatrix  $IM$ , einzuführen. Die erstgenannte stellt einen Tiefpass dar, der die Oberschwingungen (die hohen polynomialen Ordnungen) reduziert, während die zweitgenannte eine Koordinatentransformation zwischen verschiedenen Interpolationsknotensätzen durchführt, um z.B. eine Funktion in der Ordnung 4 in eine der Ordnung 8 umzurechnen. Die Filtermatrix ist definiert durch (siehe [16]):

$$FI = V \Lambda V^{-1} \quad (3.1.39a)$$

$$\Lambda = \sigma \left( \frac{N_1 + N_2 + N_3}{N} \right), \quad (N_1, N_2, N_3) > 0, \quad N_1 + N_2 + N_3 < N \quad (3.1.39b)$$

$$\sigma(\eta) = \begin{cases} 1, & 0 \leq \eta \leq \eta_c = N_c/N \\ \exp(-\alpha(\frac{\eta-\eta_c}{1-\eta_c})^s), & \eta_c \leq \eta \leq 1 \end{cases} \quad (3.1.39c)$$

Dabei beschreibt das Tripel  $(N_1, N_2, N_3)$  die Ordnung in den drei Raumdimensionen<sup>18</sup>,  $\alpha$  die Filterstärke,  $s$  die Flankensteilheit und  $N_c$  die „Cut-Off“-Ordnung. Die Interpolationsmatrix lässt sich mit Hilfe von Gleichung 3.1.19 definieren:

$$IM = V(\vec{r}_{\text{Ziel}}) \cdot V^{-1} \quad (3.1.40)$$

Dabei ist  $V(\vec{r}_{\text{Ziel}})_{ij} = \psi_j(\vec{r}_{\text{Ziel},i})$  die VANDERMONDE-Matrix, genommen an den Zielpunkten, auf die eine Funktion abgebildet werden soll.

Damit haben wir alle Operatoren beisammen, um nun zum eigentlichen Zeitschrittverfahren überzugehen. Bei der Herleitung wurde gezeigt, dass durch geschickte Wahl von Referenzelement und Basisfunktionssatz keines der Integrale im herkömmlichen Sinne aus-

<sup>18</sup>Im 2D handelt es sich entsprechend nur um ein Paar und im 1D nur um eine Zahl.

gerechnet werden musste, sondern sich diese auf einfache Algebra zurückführen lassen. Einzig die Auswertung der JACOBI-Polynome erfordert etwas mehr Aufwand, der sich jedoch mittels effizienter Rekursionsformeln bewältigen lässt.

### 3.1.3. Zeitschrittverfahren

Nach den vorangegangenen Abschnitten existiert nun ein Gleichungssystem der Gestalt

$$\partial_t \tilde{u} = L(\tilde{u}, t) \quad (3.1.41a)$$

$$L(\tilde{u}, t) = -\vec{D}\vec{f} + \tilde{g} + M^{-1}F[\dots] \quad (3.1.41b)$$

Der Übersicht wegen wurden einige Abhängigkeiten nicht mit aufgeführt. Bei diesem System ist es gelungen, die Ortsableitungen zu eliminieren, so dass nunmehr ein gewöhnliches Differentialgleichungssystem mit der Zeit  $t$  als Variable übrig bleibt. Die Lösungsmöglichkeiten reichen dabei vom einfachen EULER-Verfahren bis hin zu komplizierteren ADAM-BASHFORTH-Verfahren ([55]). In dieser Arbeit kam ein RUNGE-KUTTA-Verfahren, genauer fünfstufiges speichereffizientes RUNGE-KUTTA-Verfahren vierter Ordnung, wie in [11, 16, 27, 45] beschrieben, zur Anwendung. Um die Thematik des Zeitschrittes zu motivieren, soll hier kurz ein deutlich einfacheres Verfahren, das explizite Euler-Verfahren, skizziert werden. Dabei wird von einer Differentialgleichung (DGL) der Form  $\dot{\tilde{u}} = L(\tilde{u}, t)$  ausgegangen. Man startet mit  $\tilde{u}(t = t_0) = \tilde{u}_0$  und errechnet den nächsten Zeitschritt  $u(t + \Delta t)$  mittels:

$$\tilde{u}(t + \Delta t) = \tilde{u}(t) + \Delta t \cdot L(\tilde{u}(t), t) \quad (3.1.42)$$

Ähnlich funktioniert auch das in dieser Arbeit verwendete Runge-Kutta-Verfahren [16].

$$p^{(0)} = \tilde{u}(t) \quad (3.1.43a)$$

$$i \in [1, \dots, 5] : \begin{cases} k^{(i)} &= a_i k^{(i-1)} + \Delta t L(p^{(i-1)}, t + c_i \Delta t) \\ p^{(i)} &= p^{(i-1)} + b_i k^{(i)} \end{cases} \quad (3.1.43b)$$

$$\tilde{u}(t + \Delta t) = p^{(5)} \quad (3.1.43c)$$

Dabei bezeichnen die  $a_i$ ,  $b_i$  und  $c_i$  die RUNGE-KUTTA-Koeffizienten, während die  $p^n$  und  $k^n$  Hilfsvariablen darstellen. Hier drängt sich nun die Frage nach dem optimalen Zeitschritt auf, denn: Ist  $\Delta t$  zu kurz, vergeudet man Rechenzeit, ist er zu lang, divergiert die Lösung. Um nun das COURANT-FRIEDRICHS-LEVI-Kriterium<sup>19</sup> (CFL) ([11, 16, 64]) zu erfüllen, das

<sup>19</sup>Das CFL-Kriterium besagt, dass das Spektrum der Eigenwerte innerhalb der Stabilitätskontur des gewählten Zeitschrittverfahrens liegen muss.

heißt, um den optimalen Zeitschritt zu berechnen, gibt es einen theoretischen und einen pragmatischen Ansatz. Ersterer errechnet den Zeitschritt anhand des Eigenwertspektrums des Operators  $L$  (siehe hierzu insbesondere die oben angegebenen Quellen). Dies ist jedoch häufig nur für eher überschaubare Probleme möglich, da auch das Bestimmen von  $N_e \cdot N_p$  Eigenwerten eine gewisse Zeit in Anspruch nimmt<sup>20</sup>. Deshalb wird hier ein pragmatischer Ansatz gewählt. Dazu vergegenwärtigt man sich die physikalische Interpretation des Zeitschrittes. Dieser muss so bemessen sein, dass die Information von einem Knoten zum nächsten Knoten gelangen kann. Das heißt (mit  $c$  die charakteristische Geschwindigkeit und  $d_{min}$  der kleinste im System vorkommende Knotenabstand):

$$\Delta t \approx \frac{d_{min}}{c} \quad (3.1.44)$$

Da auch diese Berechnung einiges an Zeit erfordert, geht man vom kleinsten Knotenabstand auf den Innenkreisradius (von Dreieck, beziehungsweise Tetraeder) über. Dieser lässt sich näherungsweise mittels Volumen  $V$  und Oberfläche  $A$  bestimmen:

$$r_{\text{Innenkreis}}^k \approx \frac{V^k}{A^k} \quad (3.1.45)$$

Dieser Quotient lässt sich nun wiederum mit der Oberflächen- und Volumen-JACOBI-Determinanten approximieren, so dass sich der Zeitschritt mit:

$$\Delta t \approx c_{CFL} \frac{r_{LGL}}{c} \min_{k,e} \frac{J^{k,e}}{J^k} \quad (3.1.46)$$

ermitteln lässt. Hier wurden noch  $r_{LGL}$  als der kleinste Abstand der LEGENDRE-GAUSS-LOBATTO-Punkte im Referenzelement und  $c_{cfl}$  als eine Konstante in der Größenordnung  $c_{cfl} \approx 1$  eingeführt. Der Umweg über die JACOBI-Determinanten statt der direkten Berechnung mittels der Volumen und Oberflächen hat den Vorteil, dass bereits etwaige Anpassungen bezüglich der in Kapitel 4 behandelten gekrümmten Elementen bereits berücksichtigt sind.

### Bemerkung zur Rechenzeit

Häufig stellt sich die Frage nach der typischen Rechenzeit  $t_{CPU}$ , die für ein Projekt benötigt wird. Diese lässt sich nur sehr schwer beantworten, da sie immer vom konkreten Projekt, der verwendeten Rechentechnik und der Implementation abhängt. Die verschie-

<sup>20</sup>Nimmt man wieder ein Problem mit 20000 Elementen mit je 35 Basisfunktionen, so sind es immerhin 7 Millionen Eigenwerte.

denen methodenbedingten Einflüsse hängen dabei von der Elementanzahl, der kleinsten räumlichen Diskretisierung und der „Güte“ des physikalischen Systems ab und lassen sich dabei in folgender (vereinfachter) Formel für die Rechenzeit  $t_{CPU}$  zusammenfassen:

$$t_{CPU} = t_{\Omega_h^k} N_t (N_e + c_+) \quad (3.1.47)$$

Dabei ist  $t_{\Omega_h^k}$  die benötigte Rechenzeit je Element, wobei hier davon ausgegangen wird, dass alle Elemente gleicher Art sind, und sich so die Rechenzeit je Element als Konstante ausdrücken lässt<sup>21</sup>. Des Weiteren bezeichnet  $N_t$ , die Anzahl der Zeitschritte die notwendig ist, um ein Projekt zu berechnen. Diese Anzahl ist mit  $t_{phy}$  die physikalische Zeitspanne, die betrachtet werden soll<sup>22</sup>:

$$N_t = \frac{t_{phy}}{\Delta t} \quad (3.1.48)$$

Die zu betrachtende Zeitspanne  $t_{phy}$  hängt dabei wesentlich vom physikalischen System insbesondere der Dämpfung bzw. Güte der etwaigen Resonanzen ab, während  $\Delta t$  proportional zum kleinsten Element und damit zur Auflösung der räumlichen Diskretisierung ist. Die noch verbleibende Konstante  $c_+$  beschreibt weitere Recheneinheiten, die nicht direkt zur DG-Methode gehören, aber zum Gelingen der Simulation notwendig sind und regelmässig anfallen. Diese umfassen zum Beispiel die Frequenzfilter für die „On-The-Fly“-Fouriertransformation, Hilfsdifferentialgleichungen zur Beschreibung der Materialien und des PMLs, Feldinjektion über die TF/SF-Kontur und ähnliche Zusatzmodule.

Diese kurze Ausführung zeigt, dass die konkrete Vorhersage der Rechenzeit nicht trivial ist, und verdeutlicht, warum unter Umständen ähnliche Systeme ganz verschiedene Rechenzeiten aufweisen können.

Damit ist die kurze Einführung in die unstetige GALERKIN-Methode beendet. Jetzt folgen noch die Besonderheiten der speziellen PDGL-Systeme der MAXWELL-Gleichung und der EULER-Gleichung.

<sup>21</sup>Diese Annahme trifft insbesondere für die in dieser Arbeit behandelten runden Elemente nicht mehr zu, weswegen diese Formel in Kapitel 4 eine Erweiterung erfährt.

<sup>22</sup>Selbstverständlich machen nur ganzzahlige Zeitschritte Sinn, so dass in der Praxis immer auf die nächsthöhere ganze Zahl gerundet werden muss.

### 3.2. MAXWELL-Gleichung in DG

Nachdem im vorangegangenen Abschnitt die DG-Methode im Allgemeinen behandelt wurde, soll nun die Anwendung dieses Verfahrens auf die MAXWELL-Gleichungen diskutiert werden. Dabei dienen insbesondere die Referenzen [7, 15, 16, 20, 35, 46] als Grundlage. Da die DG-Methode hier in Erhaltungssatzform eingeführt wurde, müssen zunächst die MAXWELL-Gleichungen 2.1.1 geeignet umgeschrieben werden<sup>23</sup>:

$$\hat{Q} \partial_t \hat{u} + \vec{\nabla} \hat{f} = \hat{g} \quad (3.2.1a)$$

$$\hat{Q} = \begin{bmatrix} \epsilon(\vec{r}) & 0 \\ 0 & \mu(\vec{r}) \end{bmatrix} \quad \hat{u} = \begin{bmatrix} \vec{E}(\vec{r}, t) \\ \vec{H}(\vec{r}, t) \end{bmatrix} \quad \hat{g} = \begin{bmatrix} -\vec{j}(\vec{r}, t) \\ 0 \end{bmatrix} \quad (3.2.1b)$$

$$\hat{f} = \begin{bmatrix} \hat{f}_E \\ \hat{f}_H \end{bmatrix} = \begin{bmatrix} -\hat{e} \times \vec{H}(\vec{r}, t) \\ \hat{e} \times \vec{E}(\vec{r}, t) \end{bmatrix} \quad \hat{e} = \begin{bmatrix} \vec{e}_x \\ \vec{e}_y \\ \vec{e}_z \end{bmatrix} \quad (3.2.1c)$$

Dabei bezeichnet  $\hat{e}$  den Vektor der Einheitsvektoren und  $\vec{j} = \partial_t \vec{P}$  den Polarisationsstrom<sup>24</sup>. Zu beachten ist nun lediglich, dass es sich bei  $\hat{u}$  um einen 6-dimensionalen Vektor (für die 6 Feldkomponenten) und bei  $\hat{f}$  um einen  $6 \times 3$  - dimensionalen Vektor (für die 6 Felder sowie die 3 Raumdimensionen) handelt. Dass es sich bei dieser Schreibweise tatsächlich um die MAXWELL-Gleichungen handelt, verdeutlicht am besten ein Beispiel:

$$\epsilon \partial_t E_x = (\vec{\nabla} \times \vec{H})_x = \partial_y H_z - \partial_z H_y = \vec{\nabla} \cdot \begin{bmatrix} 0 \\ H_z \\ -H_y \end{bmatrix} = \vec{\nabla} \cdot \vec{f}_{Ex}$$

Wobei hier nur die  $E_x$ -Komponente betrachtet wurde, für die anderen Feldkomponenten gilt dieser Ansatz entsprechend. Als nächstes folgt wieder die Zerlegung der Felder in die

<sup>23</sup>Es soll an dieser Stelle nicht verwundern, dass mathematische Vektoren  $\hat{X}$  und physikalische Vektoren  $\vec{X}$  miteinander verrechnet werden. Dies dient lediglich einer übersichtlicheren Schreibweise.

<sup>24</sup>Mittels des Polarisationsstromes werden die zur Simulation dispersiver Materialien notwendigen Hilfsdifferentialgleichungen an die MAXWELL-Gleichungen angekoppelt. Dabei wirken dann diese Materialien als Quellen und Senken des elektromagnetischen Feldes. So nimmt z.B. ein LORENTZ-Oszillator, wenn er von einem Pulse getroffen wird, Energie auf, wirkt somit als Senke für das elektromagnetische Feld und gibt diese später teilweise wieder durch Strahlung ab, wirkt demnach als Quelle für das Feld.

lokalen Approximationspolynome:

$$\vec{E}(\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} \vec{E}_i^k(t) l_i^k(\vec{r}) \quad (3.2.2a)$$

$$\vec{H}(\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} \vec{H}_i^k(t) l_i^k(\vec{r}) \quad (3.2.2b)$$

$$\vec{j}(\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} \vec{j}_i^k(t) l_i^k(\vec{r}) \quad (3.2.2c)$$

Des Weiteren sollen die Materialkonstanten für jedes Element konstant<sup>25</sup> sein, das heißt:

$$\epsilon(r) = \epsilon^k = \text{konst.} \quad \mu(r) = \mu^k = \text{konst.} \quad \forall \vec{r} \in \Omega_h^k \quad (3.2.3)$$

Für eine spätere kürzere Schreibweise wird zudem noch die Impedanz  $Z$  und deren Kehrwert die Admittanz  $Y$  eingeführt.

$$Z = \frac{1}{Y} = \sqrt{\frac{\mu}{\epsilon}} \quad (3.2.4)$$

Damit ergibt sich die gewöhnliche DGL in DG-Ortsdiskretisierung zu:

$$\epsilon^k \partial_t \vec{E}^k = \vec{D}^k \times \vec{H}^k - \vec{j}^k + (M^k)^{-1} F^k \left[ \vec{n} \cdot (\hat{f}_E^k - \hat{f}_E^*) \right] \quad (3.2.5a)$$

$$\mu^k \partial_t \vec{H}^k = -\vec{D}^k \times \vec{E}^k + (M^k)^{-1} F^k \left[ \vec{n} \cdot (\hat{f}_H^k - \hat{f}_H^*) \right] \quad (3.2.5b)$$

Zur Vervollständigung des Schemas ist nur noch die Evaluation der numerischen Flussdichte notwendig. Obgleich hier die Mittelwertsflussdichte möglich wäre, welche zudem die Erhaltung der Energie sicherstellen würde, soll eine modifizierte Variante, die „Upwind Flux“, verwendet werden, der über einen freien Regelparameter  $\alpha$  verfügt. Dazu betrachtet man gleich die ganzen Integranden  $\vec{n} \cdot (\hat{f}^k - \hat{f}^*)$ :

$$\vec{n} \cdot (\hat{f}_E - \hat{f}_E^*) = \frac{1}{2\bar{Z}} \left\{ \alpha \left[ (\vec{n} \cdot \Delta \vec{E}) \cdot \vec{n} - \Delta \vec{E} \right] - \vec{n} \times Z^+ \Delta \vec{H} \right\} \quad (3.2.6a)$$

$$\vec{n} \cdot (\hat{f}_H - \hat{f}_H^*) = \frac{1}{2\bar{Y}} \left\{ \alpha \left[ (\vec{n} \cdot \Delta \vec{H}) \cdot \vec{n} - \Delta \vec{H} \right] + \vec{n} \times Y^+ \Delta \vec{E} \right\} \quad (3.2.6b)$$

<sup>25</sup>Mittlerweile wurden auch Artikel veröffentlicht, die sich mit anisotropen Materialkonstanten beschäftigen [28]. Da solche in dieser Arbeit nicht vorkommen und gleichzeitig eine etwas kompliziertere numerische Flussdichte erfordern, soll hier nicht weiter darauf eingegangen werden.

Dabei wurden folgende Abkürzungen (mit  $X^-$  der Funktionswert von Element  $\Omega_h^k$  und  $X^+$  der Funktionswert des jeweiligen Nachbarelementes) verwendet:

$$\Delta X = X^- - X^+ \qquad \bar{X} = \frac{1}{2}(X^- + X^+) \qquad (3.2.7a)$$

Die Gleichung 3.2.6 besteht im Wesentlichen aus zwei Beiträgen. Der erste Teil:

$$-\frac{Z^+}{2\bar{Z}}\vec{n} \times \Delta\vec{H} \qquad \text{beziehungsweise} \qquad \frac{Y^+}{2\bar{Y}}\vec{n} \times \Delta\vec{E}$$

ist nichts anderes, als die bereits bekannte Mittelwertsflussdichte, während der zweite Teil:

$$\frac{1}{2\bar{Z}} \left[ (\vec{n} \cdot \Delta\vec{E}) \cdot \vec{n} - \Delta\vec{E} \right] \qquad \text{beziehungsweise} \qquad \frac{1}{2\bar{Y}} \left[ (\vec{n} \cdot \Delta\vec{H}) \cdot \vec{n} - \Delta\vec{H} \right]$$

den Sprung in der Tangentialkomponente von  $\vec{E}$  beziehungsweise  $\vec{H}$  über die Elementgrenzen berücksichtigt. Jetzt wird auch die Bedeutung des Parameters  $\alpha \in [0, 1]$  klar, mit dem zwischen dem „Central Flux“ ( $\alpha = 0$ ) und dem klassischen „Upwind Flux“ ( $\alpha = 1$ ) gewichtet werden kann. Ist  $\alpha = 0$ , so hat man volle Energieerhaltung, kauft sich unter Umständen unerwünschte Moden der Felder beziehungsweise zum Teil deutlich kürzere Zeitschritte ein. Wählt man hingegen  $\alpha = 1$ , so werden diese Sprünge zwar minimiert, man dämpft dabei aber das System, so dass eine zusätzliche Absorption die Energieerhaltung verletzt. In der Praxis hat sich jedoch gezeigt, dass der „kontinuierlicheren“ Lösung der Vorzug zu geben ist, da diese eher der physikalischen Realität entspricht, so dass sich die oben beschriebene Dämpfung hauptsächlich auf die unerwünschten, nichtphysikalischen Lösungen auswirkt. Damit kann schließlich die Stabilität des Systems erhöht werden [11].

### 3.2.1. Randbedingungen

Nachdem im bisherigen Verlauf lediglich die DG-Methode auf die Lösung im Volumen des Simulationsgebiets  $\Omega_h$  behandelt wurde, soll es hier um die noch benötigten Randbedingungen gehen. Diese lassen sich in zwei wichtige Bereiche aufteilen, die Anfangswerte für die Simulation und die „eigentlichen“ Randbedingungen auf dem Rand  $\partial\Omega_h$  des betrachteten Volumenstückchens. Während die Anfangswerte stark von der gewählten Aufgabenstellung abhängen, lassen sich die klassischen Randbedingung etwas allgemeiner klassifizieren. Zu diesen Bedingungen gehören insbesondere der perfekte elektrische Leiter (engl.: Perfect Electric Conductor: PEC<sup>26</sup>), der perfekte magnetische Leiter (engl.: Perfect Ma-

<sup>26</sup>Beim PEC verschwinden die tangentialen  $\vec{E}$ -Feld-Komponenten. Er wird auch dann verwendet, wenn ein Streuer aus einem perfekten Metall simuliert werden soll.



gnetic Conductor: PMC<sup>27</sup>) sowie die sogenannte absorbierende Randbedingung (Silver-Müller Perfectly Matched Layer: SM-PML). Letztere hat insbesondere die Aufgabe, die aus dem Simulationsgebiet hinauslaufenden elektromagnetischen Wellen zu absorbieren. Diese Randbedingung ist aber nur für Wellenvektoren  $\vec{k}$  parallel zur Oberflächennormalen exakt. Die ersten beiden Randbedingungen beschreiben hingegen perfekte elektrische beziehungsweise magnetische Spiegel. Sie dienen vor allem zur Berechnung von Kavitäten oder zur Reduzierung des Rechenaufwandes, wenn Symmetrieebenen berücksichtigt werden können. Die Randbedingungen werden über die Felddifferenzen (siehe hierzu Gleichung 3.2.6) bei der Flussberechnung im Gleichungssystem berücksichtigt. Daher genügt hier die Angabe der zu wählenden Differenzen  $\Delta\vec{E}$  und  $\Delta\vec{H}$  ([16, 45, 46]):

**Tabelle 3.2.1.:** Randbedingungen für die MAXWELL-Gleichungen

	$\Delta\vec{E}$	$\Delta\vec{H}$
PEC	$2\vec{E}^-$	0
PMC	0	$2\vec{H}^-$
SM-PML	$2\vec{E}^-$	$2\vec{H}^-$

### 3.2.2. Implementationstests

Nach all der Theorie folgen nun die ersten echten Simulationen. Bevor es allerdings tatsächlich an die Nanooptik geht, muss die Implementation des Codes getestet sowie die noch freien Parameter bestimmt werden. Zu diesen gehören der kritische Zeitschritt  $\Delta t$  und die PML-Parameter. Als Tests der Implementation stehen die reinen MAXWELL-Gleichungen, Randbedingungen, TF/SF-Kontur, Materialmodelle und Frequenzfilter auf dem Programm.

#### Bestimmung des kritischen Zeitschrittes

Als sicherlich wichtigste Aufgabe folgt nun die Bestimmung des Zeitschrittes  $\Delta t$ . Als Ansatz für diesen Parameter gilt nach Gleichung 3.1.46, die hier noch mal wiederholt werden soll:

$$\Delta t \approx c_{CFL} \frac{r_{LGL}}{c} \min_{k,e} \frac{J^{k,e}}{J^k}$$

<sup>27</sup>Beim PMC verschwinden die tangentialen  $\vec{H}$ -Feld-Komponenten.

Dabei ist die kritische Geschwindigkeit  $c$  gleich der Vakuumlichtgeschwindigkeit  $c_0$  als größter möglicher Geschwindigkeit einer elektromagnetischen Welle. Es bleibt noch die Bestimmung der CFL-Variablen  $c_{CFL}$ . Dazu wurde für jede der drei Dimensionen die Zeitentwicklung der Grundmode in einer Kavität als Funktion der Ortsdiskretisierung  $h$ , der Ordnung der Approximationspolynome  $N$  und des  $c_{CFL}$ -Parameters berechnet. Hier ist  $h$  ein Maß für die Größe der Elemente, und das Inverse von  $h$  ein Maß für die Anzahl der Elemente je Einheitslänge. Damit entspricht  $1/h$  im 1D gleichzeitig der Anzahl der Elemente  $N_e$ . Als Kavitäten kamen dabei die Einheitswürfel in den drei Dimensionen, wie in Tabelle 3.2.2 definiert, zum Einsatz: Als Startwert wurde die Grundmode ( $n = 1$ ,

**Tabelle 3.2.2.:** Einheitswürfel  $\Omega_h$  in verschiedenen Dimensionen, inklusive der Lösung für die  $E_z$ -Komponente mit der zugehörigen Frequenz  $\omega$ .

	$\Omega_h$	$E_z$	$\omega$
1D	$\vec{r} \in [0, 1]$	$\sin(n\pi x)$	$\pi n$
2D	$\vec{r} \in [0, 1] \times [0, 1]$	$\sin(n\pi x) \sin(m\pi y)$	$\pi \sqrt{n^2 + m^2}$
3D	$\vec{r} \in [0, 1] \times [0, 1] \times [0, 1]$	$\sin(n\pi x) \sin(m\pi y)$	$\pi \sqrt{n^2 + m^2}$

**Tabelle 3.2.3.:** Verwendete Diskretisierungen  $h$  in Abhängigkeit der Ordnung  $N$  und der Dimension zur Bestimmung von  $c_{CFL}$ .

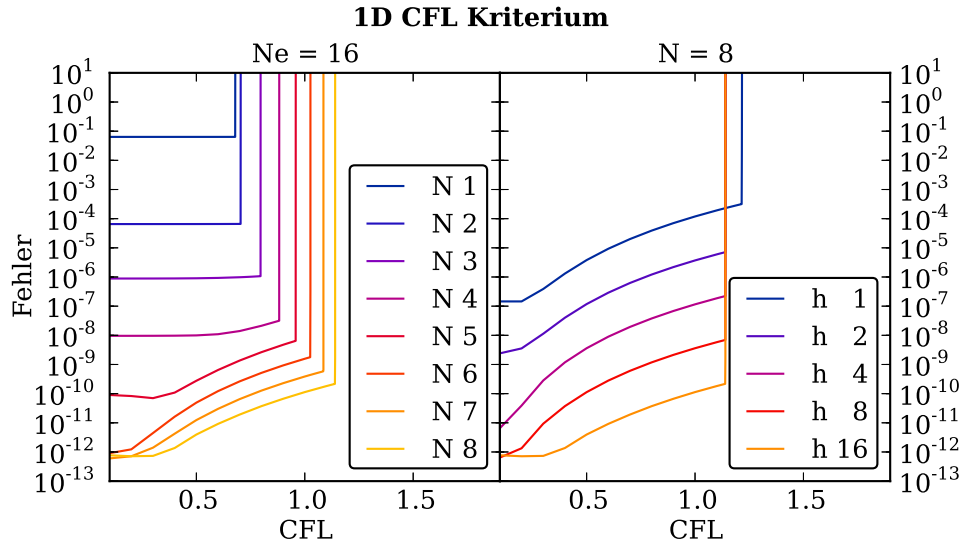
N	$\frac{1}{h}$ : 1D & 2D	$\frac{1}{h}$ : 3D
1 - 2	4 - 64	4 - 32
3 - 4	2 - 32	2 - 16
5 - 8	1 - 16	1 - 8

$m = 1$ ), wie in Tabelle 3.2.2 aufgeführt, angesetzt. Danach wurde nach 100 Periodendauern  $T = 2\pi/\omega$  der Fehler der  $E_z$  Komponente:

$$Fehler = \max_{N_p, N_e} |E_z^{analytisch} - E_z^{DG}| \quad (3.2.8)$$

im Vergleich zur analytischen Lösung ausgewertet. Diese Ergebnisse sind exemplarisch für ausgewählte Parameter in Abbildung 3.2.1 zusammengestellt. Deutlich sieht man, dass der kritische  $c_{CFL}$ -Parameter mit steigender Ordnung  $N$  zunimmt (linke Seite), jedoch für eine gewählte Ordnung „relativ“ unabhängig von der konkreten Diskretisierung ist<sup>28</sup>.

<sup>28</sup>Der Ausreißer für  $h = 1$  geht auf Rundungsfehler zurück, da, um exakt 100 Perioden rechnen zu können, der Zeitschritt so angepasst wird, dass  $\Delta t$  ein ganzzahliger Teil von  $100T$  ist, und in diesem Zusam-



**Abbildung 3.2.1.:** Bestimmung des CFL-Zeitschrittes im 1D. Links für ein Netz mit 16 Elementen und verschiedenen Ordnungen. Rechts für die achte Ordnung und verschiedenen Diskretisierungen. Dabei ist jeweils der Fehler der Simulation nach 100 Periodendauern in Abhängigkeit der  $c_{CFL}$  Konstanten dargestellt. Deutlich sieht man den kritischen Zeitschritt, bei dem der Fehler sehr stark wächst. Zudem ist die Abhängigkeit der  $c_{CFL}$  Konstante von der Ordnung und deren Unabhängigkeit von der Diskretisierung sichtbar.

Die so aus allen Rechnungen bestimmten „kleinsten“ Parameter  $c_{CFL}(N)$  sind in Tabelle 3.2.4 zusammengetragen. Weiterhin bietet die Abbildung die Möglichkeit, den Verlauf des Fehlers auch unterhalb des kritischen Zeitschrittes zu studieren. Hier fällt auf, dass für die niedrigen Ordnungen  $N < 4$ , Fehler  $> 10^{-7}$  der Fehler unabhängig vom konkreten  $c_{CFL}$ , und damit vom Zeitschritt ist (solange  $\Delta t < \Delta t^{krit.}$ ). Dies deutet darauf hin, dass der Fehler durch die räumliche Diskretisierung bestimmt ist. Anders verhält es sich für die hohen Ordnungen und hier speziell  $N > 5$ . Hier lässt sich insbesondere für  $c_{CFL} < 0.3$ , trotz höherer Ordnung, keine Steigerung der Genauigkeit erreichen. Dies wird besonders im Vergleich von  $N = 8$ ,  $h = 8$  mit  $N = 8$ ,  $h = 16$  deutlich, die, obwohl Letztere die doppelte räumliche Auflösung besitzt, den gleichen Fehler aufweisen. Ursache ist hier die endliche Zahlendarstellung im Computer, welche eine Präzision von ca. 12 - 14 Digits erlaubt. Im Zwischenbereich von  $0.3 < c_{CFL} < c_{CFL}^{krit.}$  steigt der Fehler recht schnell von  $10^{-12}$  um, je nach Ordnung  $N$ , 2 bis 4 Größenordnungen. In diesem Bereich bilden die endliche Stabilität der Matrizen und die durch den „Upwind Flux“ eingeführte künstliche Dämpfung die Hauptfehlerquelle. Soweit nicht anders vermerkt, wurde in allen folgenden

---

menhang immer die im Vergleich zur Zeitschrittschätzung nächsthöhere Zeitschrittanzahl verwendet wurde.

**Tabelle 3.2.4.:** CFL Konstanten, in Abhängigkeit der Ordnung  $N$  und der Dimension.

N	1D	2D	3D
1	0.678	0.477	0.33
2	0.704	0.540	0.42
3	0.796	0.665	0.54
4	0.882	0.745	0.64
5	0.958	0.834	0.71
6	1.025	0.878	0.77
7	1.084	0.936	0.82
8	1.139	0.968	0.86

Rechnungen  $c_{CFL} = 0.8c_{CFL}^{krit.}$  verwendet.

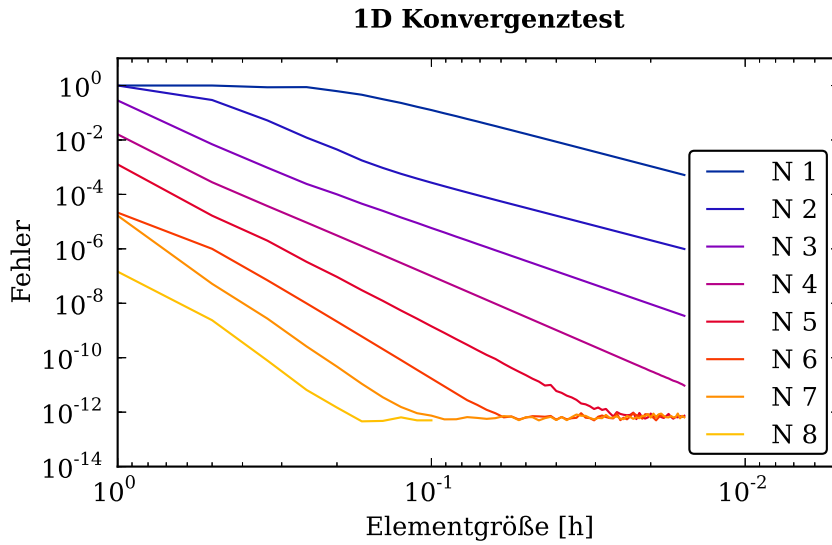
### Konvergenzanalyse

Nachdem der kritische Parameter  $c_{CFL}^{krit.}$  zur Schätzung des optimalen Zeitschrittes bestimmt wurde, kann nun die eigentliche Konvergenzanalyse folgen. Ohne auf weitere Details einzugehen<sup>29</sup>, sei an dieser Stelle vermerkt, dass bei Verwendung des „Upwind Flux“ für die MAXWELL-Gleichung in der, in dieser Arbeit, verwendeten Form folgende scharfe Fehlerbegrenzung gilt:

$$|\hat{u}^{exakt} - \hat{u}^{DG}| \leq (c_1 + c_2 \cdot T) \cdot h^{N+1} \quad (3.2.9)$$

Dabei ist  $\hat{u}$  die Lösung der PDGL,  $c_1$  und  $c_2$  systemabhängige Konstanten,  $h$  wieder ein Maß für die charakteristische Größe des größten Elementes sowie  $N$  die Ordnung der Approximationspolynome. Da die Ordnung in der Literatur häufig mit  $p$  statt mit  $N$  abgekürzt wird, spricht man auch von  $hp$ -Konvergenz. Diese besagt, dass der Fehler umso kleiner wird, je höher die Ordnung beziehungsweise je kleiner die Elemente sind. Des Weiteren besagt diese Fehlergrenze, dass der Fehler im Wesentlichen linear mit der Zeit steigt. Wenn die der Arbeit zugrunde liegende Implementation richtig ist, so muss sie die gleiche Konvergenzrate von  $Fehler \approx h^{N+1}$  zeigen. Dazu wurden für die höheren Dimensionen (2D & 3D) die bereits vorhanden Ergebnisse von oben mit  $c_{CFL} = 0.1$  ausgewertet (hier nicht gezeigt), während für den 1D-Fall die in Abbildung 3.2.2 dargestellten Ergebnisse mit  $c_{CFL} = 0.2c_{CFL}^{krit.}$  und deutlich mehr Zwischenwerten für  $h = 1/N_e$  ermittelt wurden. Dabei stellte sich heraus, dass die Ordnungen  $N = 2...8$  die erwartete Konvergenzrate zeigen. Eine Ausnahme bildet die Ordnung  $N = 1$ , die eine ähnliche Konvergenzrate wie  $N = 2$  zeigt. Eine Besonderheit ist noch die Ordnung  $N = 8$ , für die relativ wenige Daten

<sup>29</sup>Eine eingehende Untersuchung zu diesem Thema findet sich in [15, 16].



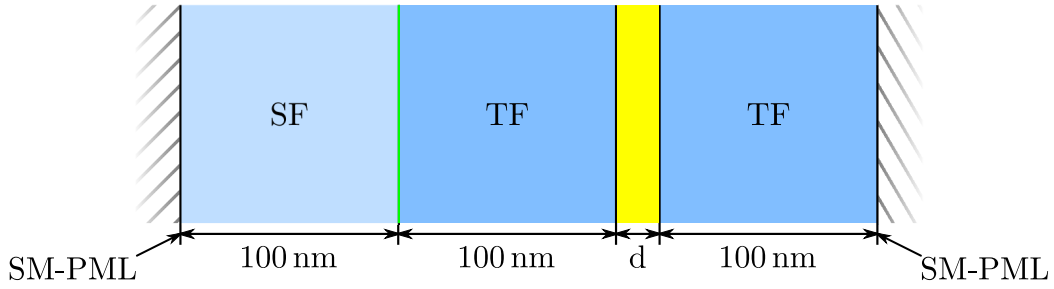
**Abbildung 3.2.2.:** Konvergenzanalyse für lineare Elemente im 1D. Dargestellt ist der Fehler der Simulation der Einheitskavität im 1D als Funktion der Ordnung und der Elementgröße  $h$ . Dabei ist die Abnahme des Fehlers um so größer, je höher die Ordnung ist. Des Weiteren sieht man sehr schön die Fehleruntergrenze bei  $12^{-12}$  als Folge der finiten Zahlendarstellung im Computer.

zur Verfügung stehen, um die Konvergenzrate über einen größeren Dynamikumfang von  $h$  bestimmen zu können, da hier bereits für  $N_e = 6$  die systembedingte Untergrenze des Fehlers erreicht wird. Des Weiteren macht sich auch hier wieder die endliche Genauigkeit der Zahlendarstellung bemerkbar. Nachdem der niedrigste Fehler erreicht ist, steigt dieser mit kleiner werdendem  $h$  sogar wieder leicht an. Ursache ist wieder die endliche Präzision und die damit verbundenen Rundungsfehler.

## Materialanalyse

Nach den etwas technischen Analysen geht es nun um die Überprüfung der Implementation der Materialmodelle, des DFT-Filters sowie der Feldinjektion mittels TF/SF-Kontur. Diese Rechnungen lassen sich gut im 1D durchführen<sup>30</sup>, da hier alle Randbedingungen exakt erfüllt werden können und man sich so noch nicht um den PML kümmern muss. Darüber hinaus sind alle weiteren Gleichungen und Ansätze identisch zu den höheren Dimensionen. Mit der Wechselwirkung des elektromagnetischen Feldes mit einer planparallelen Platte steht zudem ein etwas interessanteres Testsystem, als es die „einfachen“ Kavitäten waren,

<sup>30</sup>Genau genommen unterscheidet der Code nicht zwischen 1D, 2D und 3D. Es werden immer alle 6 Feldkomponenten mit denselben Klassen und Funktionen berechnet. Lediglich die notwendigen Operatoren werden, wie in Abschnitt 3.1.2 dargestellt, an die entsprechende Dimension angepasst.



**Abbildung 3.2.3.:** 1D Planparallele Platte: Das System besteht aus vier Bereichen (SF, TF, planparallele Platte der Dicke  $d$ , TF) und wird durch die im 1D exakte SM-PML Randbedingung abgeschlossen. An der Grenze zwischen dem SF-Bereich und dem TF-Bereich wird die Anregung in das System injiziert.

zur Verfügung, dessen Lösung auch analytisch zugänglich ist [4]. Das untersuchte System besteht aus einer planparallelen Platte der Dicke  $d$  aus den Materialien Glas ( $n = 1.5$ ,  $d = 1 \mu\text{m}$ ), DRUDE-Silber ( $d = 0.01 \mu\text{m}$ ) und DRUDE-LORENTZ-Gold ( $d = 0.02 \mu\text{m}$ ). Dabei entsprechen die Materialparameter der Metalle denen aus Tabelle 2.3.1. Eine Skizze ist in Abbildung 3.2.3 dargestellt. Bestimmt wird hierbei Reflexion, Transmission und Absorption im Wellenlängenbereich von  $0.4 - 2.0 \mu\text{m}$ . Dazu wurde ein differenzierter GAUSS-Puls mit der Mittenwellenlänge von  $0.7 \mu\text{m}$  verwendet. Während der Simulation wurde das  $\vec{E}$ - und  $\vec{H}$ -Feld an 100 Frequenzpunkten im oben genannten Wellenlängenbereich gefiltert. Die Ortsdiskretisierung der planparallelen Platte bestand aus Elementen 4. Ordnung der Länge  $0.1 \mu\text{m}$ ,  $0.01 \mu\text{m}$  und  $0.02 \mu\text{m}$  für Glas, Silber beziehungsweise Gold. Die analytischen Reflexions- und Transmissionskoeffizienten erhält man mittels [4]:

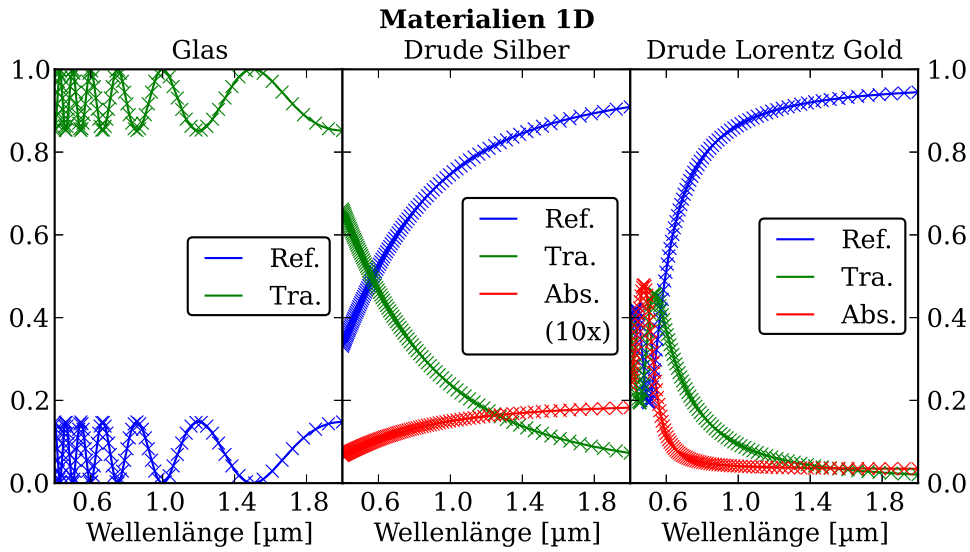
$$\check{m} = \frac{\check{n}_2}{n_1} \quad (3.2.10a)$$

$$\check{r} = \frac{1 - \check{m}}{1 + \check{m}} \quad (3.2.10b)$$

$$\check{r}_{Slab} = \frac{\check{r}(1 - e^{\check{r}^2 \frac{\omega}{c_0} \check{n}_2 d})}{1 - \check{r}^2 e^{\check{r}^2 \frac{\omega}{c_0} \check{n}_2 d}} \quad (3.2.10c)$$

$$\check{t}_{Slab} = \frac{4\check{m}}{(\check{m} + 1)^2} \frac{e^{-\check{r} \frac{\omega}{c_0} n_1 d}}{(e^{-\check{r} \frac{\omega}{c_0} \check{n}_2 d} - \check{r}^2 e^{\check{r} \frac{\omega}{c_0} \check{n}_2 d})} \quad (3.2.10d)$$

Dabei ist  $\check{n}_1$  der komplexe Brechungsindex der planparallelen Platte und  $n_1$  der Brechungsindex des umgebenden Mediums (in diesem Fall Luft mit  $n_1 = 1$ ). Der Vergleich zwischen analytischer Lösung und der DG-Simulation ist in Abbildung 3.2.4 für die drei verschiedenen Fälle dargestellt. Als Wichtigstes fällt die sehr gute Übereinstimmung zwischen Theorie



**Abbildung 3.2.4.:** Vergleich der Transmissions- und Reflektionsspektren für verschiedene Materialmodelle. Linien: Analytische Lösung, Kreuze: DG-Lösung. Bei den Werten von Silber wurde die Absorption um den Faktor Zehn vergrößert gezeichnet.

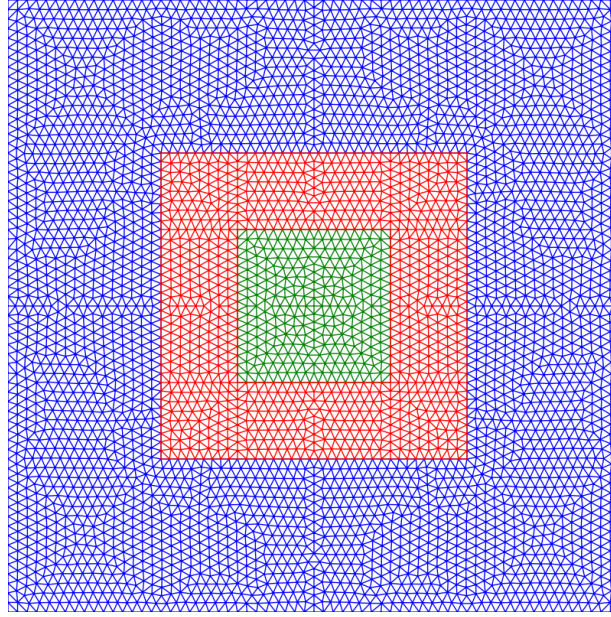
und (numerischem) Experiment auf, so dass Pulsinjektion, DFT-Filter, Materialmodelle und Randbedingung gut zu funktionieren scheinen.

### PML-Parameter

Nachdem nun wesentliche Bestandteile des Codes verifiziert wurden, bleibt noch der optimale Parametersatz für die PML zu bestimmen. Vereinfachend wird in der Herleitung wieder nur die Absorption in  $x$ -Richtung behandelt. Die anderen Raumrichtungen ergeben sich wieder sinngemäß. Als erste Eigenschaft wird angenommen, dass das Absorptionsprofil des PMLs einem polynomialen Verlauf folgt, um den Übergang zwischen der eigentlichen Rechendomäne und dem Absorptionsbereich so sanft wie möglich zu gestalten. Mit  $l$  als Abstand von der Rechendomäne-PML-Grenze in  $x$ -Richtung,  $\Delta l$  als PML-Dicke,  $S$  als Grad des Polynoms und  $R$  als Absorptionsstärke des PMLs folgt für das Absorptionsprofil ([2, 7, 20, 35, 46]):

$$s_x = R \cdot \frac{S+1}{2\Delta l} \cdot \left( \frac{l}{\Delta l} \right)^S \quad (3.2.11)$$

Es sind noch geeignete Parameter  $R$  und  $S$  zu ermitteln. Dazu wurden im 2D zwei Systeme gerechnet, ein Testsystem mit einer von einem PML umschlossenen Luftdomäne sowie ein deutlich größeres Referenzsystem ohne PML. Beide sind in Abbildung 3.2.5 dargestellt. Dabei besteht das Testsystem aus dem grünen (Luft) und roten (PML) Bereich und endet am Außenrand der roten Fläche. Das Referenzsystem setzt sich hingegen über alle Bereiche



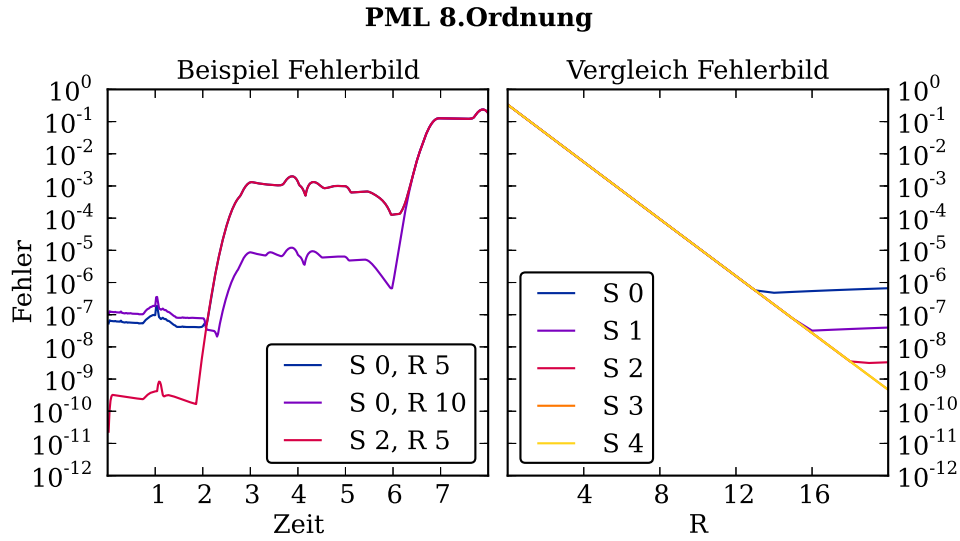
**Abbildung 3.2.5.:** Netz zum Berechnen des PML Fehlers. Grün: Luft, rot: PML und blau: zusätzliche Referenzdomäne. Größe:  $6 \times 6$  Längeneinheiten.

(hier alle samt Luft) fort. Abgeschlossen werden die Systeme jeweils (das Testsystem nach dem roten, das Referenzsystem nach dem blauen Bereich) mit den PEC-Randbedingungen. Diese führen zu einer starken Reflexion, so dass nicht absorbierte Wellen leicht sichtbar werden. Als Startwert diene eine in der Mitte zentrierte GAUSSsche Glocke in  $E_z$  mit  $\sigma = 0.2$ :

$$E_z = e^{\frac{1}{2} \frac{x^2+y^2}{\sigma^2}} \quad (3.2.12)$$

und die Ordnung der Approximationspolynome lag bei  $N = 8$ . Nun wurden der Feldfehler  $E_z$  in der inneren (grünen) Domäne als Differenz zwischen Referenzsystem und Testsystem in Abhängigkeit von der Zeit  $t$  und dem Parametern  $R$  und  $S$  bestimmt. Das Ergebnis ist in Abbildung 3.2.6 dargestellt. Dabei zeigt die linke Figur den Zeitverlauf der maximalen Felddifferenzen für ausgewählte Parameter  $R$  und  $S$ . Hier fallen im Wesentlichen drei wichtige Zeiten auf. Erstens trifft bei  $t = 1$  das Maximum des auseinander laufenden Pulses die PML-Region, zweitens erreichen die am PEC am Ende der PML-Region zurückreflektierten Wellen bei  $t = 2$  wieder die innere Domäne, und drittens gelangen bei  $t = 6$  die am PEC des Referenzsystems reflektierten Wellen wieder in den grünen Bereich. Daher ist für die Bestimmung der Güte des PMLs nur der Zeitbereich  $0 \leq t < 6$  entscheidend, in der das Referenzsystem ohne Einflüsse vom Rand läuft. Genauer wurde in dieser Arbeit das Maximum des Fehlers im Bereich  $0 \leq t \leq 4$  als Maß für die Qualität des PMLs verwendet. Anhand der Abbildung 3.2.6 links lässt sich auch gut die Wirkung der Parameter  $S$  und  $R$  studieren. So sorgt eine hohe Ordnung  $S$  des Absorptionsprofils tatsächlich für einen





**Abbildung 3.2.6.:** Optimierung der PML Parameter. Im ersten Diagramm ist das Zeitverhalten des PML-Fehlers für verschiedene Parameter gezeigt. Im zweiten Diagramm ist der maximale PML-Fehler für  $t < 4$  für die verschiedenen Parameter  $S$  und  $R$  gezeigt.

kleineren Fehler im ersten Zeitfenster von  $0 \leq t \leq 2$  (vgl.  $S = 0, R = 5$  und  $S = 2, R = 5$ ), während dieser Parameter keinen Einfluss auf den zweiten Zeitbereich von  $2 \leq t \leq 4$  hat. Die Fehlerwerte für  $S = 0, R = 5$  und  $S = 2, R = 5$  zeigen in diesem Bereich exakt den gleichen Verlauf. Dieser Bereich, der durch die Absorption des PMLs bestimmt wird, lässt sich daher am besten über die PML-Stärke  $R$  regeln. So nimmt erwartungsgemäß die Absorption, und damit die Güte des PMLs, mit steigendem Wert für  $R$  zu (vgl.  $S = 0, R = 5$  und  $S = 0, R = 10$ ). Gleichzeitig steigt aber auch die erste Reflexion an der Luft-PML-Grenze minimal. Es gilt nun zwischen einer möglichst großen Absorption und einer nicht zu großen Reflexion an der Luft-PML-Grenze zu balancieren. Hier hilft die rechte Abbildung, in welcher für jeden Parameter  $S$  der maximale Fehler für  $0 \leq t \leq 4$  als Funktion der PML-Stärke  $R$  aufgetragen wurde. Zuerst sinkt der Fehler sehr schnell mit steigendem  $R$  und unabhängig von  $S$ , bevor ab  $R > 13$  die Unterschiede aufgrund der verschiedenen Reflexionen an der Luft-PML-Grenze deutlich werden. Vor Erreichen der „Sättigung“ ist der Fehler durch die Absorption des PMLs dominiert. Dass der minimale Fehler für  $S > 2$  noch nicht erreicht wurde, liegt hier an den noch zu niedrigen PML-Stärken  $R$ .

Da für die meisten Probleme eine Genauigkeit von 4 Digits ausreicht (der Fehler liegt dann bei ca. 0.01%), häufig mit deutlich weniger Elementen und Ordnungen für den PML gerechnet wird und so mit einer deutlich schlechteren Ortsauflösung sowohl das exponentiell abklingende Feld wie auch das Absorptionsprofil aufgelöst werden muss, wurde im Folgenden, soweit nicht anders erwähnt, mit  $R = 10$  und  $S = 0$  gerechnet.

### 3.3. EULER-Gleichung in DG

Nachdem der Formalismus der MAXWELL-Gleichungen vorgestellt und erste Verifikationen dazu präsentiert wurden, soll es jetzt um das zweite große PDGL-System, die EULER-Gleichungen, gehen. Dies gestaltet sich dahingehend etwas leichter, da sie per se eine natürliche Erhaltungssatzdarstellung besitzen, auf die hier zurückgegriffen werden kann:

$$\partial_t \hat{u} + \vec{\nabla} \hat{f} = \hat{g} \quad (3.3.1a)$$

$$\hat{u} = \begin{bmatrix} \rho \\ \rho \vec{v} \\ U \end{bmatrix} \quad \hat{f} = \begin{bmatrix} \vec{f}_\rho \\ \hat{f}_{\rho \vec{v}} \\ \vec{f}_U \end{bmatrix} = \begin{bmatrix} \rho \vec{v} \\ \rho \vec{v} \otimes \vec{v} + p \\ (U + p) \cdot \vec{v} \end{bmatrix} \quad \hat{g} = \begin{bmatrix} \delta \rho \\ \vec{F} \\ \delta U \end{bmatrix} \quad (3.3.1b)$$

Dabei wurde wieder ein Quellterm für die Teilchenerzeugung  $\delta \rho$ , die Kraftdichte  $\vec{F}$  (als Quelle beziehungsweise Senke der Impulsänderung) und der Quellterm für die Energieänderung  $\delta U$  eingeführt. Nun folgt wieder die Zerlegung in die Approximationspolynome<sup>31</sup>:

$$\rho(\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} \rho_i^k(t) l_i^k(\vec{r}) \quad (3.3.2a)$$

$$[\rho \vec{v}](\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} [\rho \vec{v}]_i^k(t) l_i^k(\vec{r}) \quad (3.3.2b)$$

$$U(\vec{r}, t) \approx \bigoplus_{k=1}^{N_e} \sum_{i=1}^{N_p} U_i^k(t) l_i^k(\vec{r}) \quad (3.3.2c)$$

Die Zerlegung des Quellterms  $\hat{g}$  erfolgt sinngemäß. Im Gegensatz zur MAXWELL-Gleichung kommt hier nun die schwache Form der DG-Methode zum Einsatz:

$$\partial_t \hat{u}^k(t) = (M^k)^{-1} (\vec{S}^k)^T \hat{f}^k + \hat{g}^k - (M^k)^{-1} F^k[\vec{n} \hat{f}^*] \quad (3.3.3)$$

Als numerische Flussdichte wird hier die lokale LAX-FRIEDRICHS-Flussdichte verwendet:

$$(n_x f_x + n_y f_y + n_z f_z)^* = n_x \bar{f}_x + n_y \bar{f}_y + n_z \bar{f}_z + \frac{\lambda}{2} \Delta u \quad (3.3.4)$$

<sup>31</sup>Hier gilt es jetzt zu beachten, was die Erhaltungsgrößen sind. So wird jetzt die Teilchenstromdichte  $[\rho \vec{v}]$  die Erhaltungsgröße, so dass das Produkt aus Dichte  $\rho$  und Geschwindigkeit  $\vec{v}$  diskretisiert werden muss. Gelegentlich findet man statt  $[\rho \vec{v}]$  auch  $\vec{j}$ , dann gilt es jedoch sorgfältig zwischen der Teilchenstromdichte und der mit dem gleichen Symbol verknüpften Ladungstromdichte  $q\rho \vec{v}$  (mit  $q$  als Ladung pro Teilchen) zu unterscheiden.

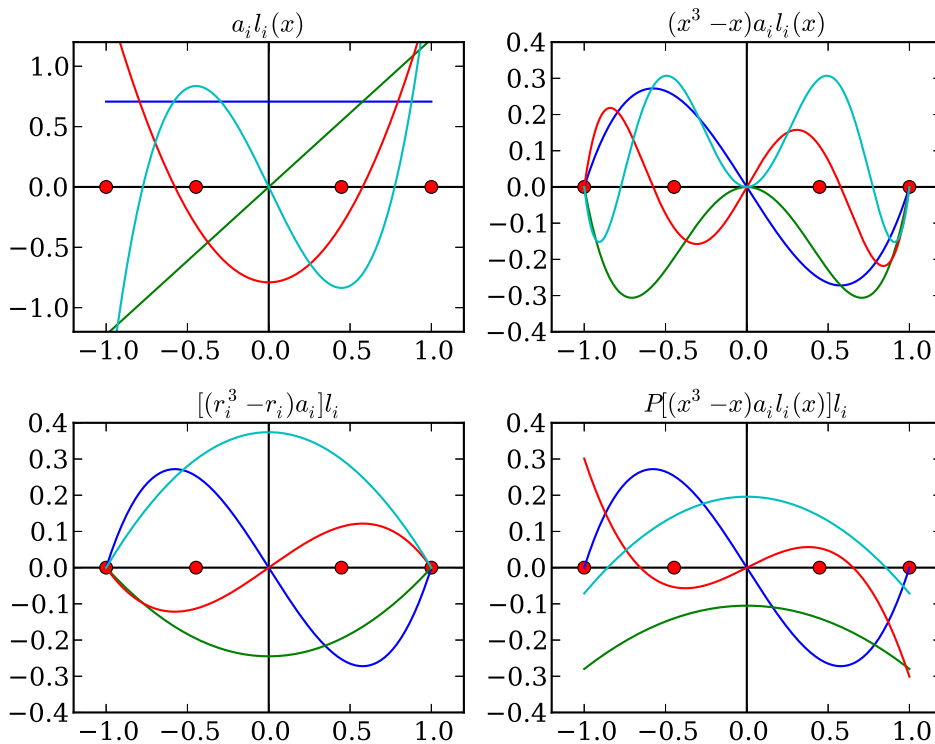
Dabei ist  $\lambda$  die approximierte lokale maximale linearisierte Schallgeschwindigkeit:

$$\lambda = \max_{\partial\Omega_h^+, \partial\Omega_h^-} \left( |\vec{v}| + \sqrt{\left| \frac{\gamma p}{\rho} \right|} \right) \quad (3.3.5)$$

Hier ist  $\gamma$  wieder der adiabatische Exponent.

### Interpolationsfehler

#### Interpolationsfehler 1D



**Abbildung 3.3.1.:** Dargestellt ist die Entstehung des Interpolationsfehlers im 1D für ein Element dritter Ordnung bei Multiplikation mit  $f(x) = x^3 - x$ . Oben Links: Ausgangsfunktion (hier die JACOBI-Polynome) bis zur dritten Ordnung. Oben Rechts: das „echte“ Produkt. Unten Links: die Bildung des Produktes mittels Evaluierung an den Stützstellen  $r_i$  mit nachfolgender Interpolation. Unten Rechts: das „echte“ Produkt der Ausgangsfunktionen, projiziert auf den Raum der Polynome bis zur Ordnung Drei. Deutlich ist der Unterschied zwischen der Auswertung des Produktes an den Stützstellen und der „echten“ Multiplikation mit anschließender Projektion zu sehen.

Im Gegensatz zu den MAXWELL-Gleichungen sind die EULER-Gleichungen nichtlinear. Des Weiteren muss man sich vergegenwärtigen, dass alle Berechnungen nicht mit den

Polynomen an sich, sondern nur noch mit den Koeffizienten und damit mit den Werten der Felder an den Interpolationsknoten durchgeführt werden. Dies ist richtig, solange der Raum der Polynome der Ordnung  $N$  nicht verlassen wird, das heißt, solange nur Linearkombinationen von Polynomen der Ordnung  $N$  vorkommen, wie dies bei den MAXWELL-Gleichungen der Fall ist. Im nichtlinearen Fall treten nun nicht nur Linearkombinationen, sondern auch Produkte und Potenzen auf (z.B. in der Berechnung des Flusses und des Druckes). Dies führt dazu, dass die Interpolation nicht mehr die korrekte Funktion darstellt, sobald die Funktion die Ordnung  $N$  verlässt. Derartige Fehler werden auch Aliasfehler genannt. Dies ist in Abbildung 3.3.1 für den Fall der Multiplikation mit  $x^3 - x$  dargestellt. Dabei stellt die obere Zeile die Ausgangsfunktionen, in diesem Beispiel die JACOBI-Polynome, in der Basis der LEGENDRE-Polynome  $a_i l_i(x)$  bis zur dritten Ordnung und die Zielfunktion  $(x^3 - x)a_i l_i(x)$  dar. In der unteren Zeile ist die Interpolation der Zielfunktion an den Punkten  $r_i$  und die Projektion  $P$  der Zielfunktion auf die Basis der LEGENDRE-Funktion bis zur Ordnung  $N = 3$  dargestellt. Ersichtlich ist der Unterschied zwischen der exakten Zielfunktion und den Approximationen zu sehen. Deutlich wird auch der Unterschied zwischen der exakten Approximation über die Projektion und der einfachen Interpolation aufgezeigt. Diese Interpolationsfehler können zu unkontrolliertem Verhalten führen. Eine Möglichkeit, sie zu umgehen, wäre, die Berechnung des Flusses in einem Raum mit höherer Interpolationsordnung (in diesem Beispiel addieren sich die Ordnung der Interpolationspolynome  $N$  und die der Multiplikationsfunktion  $M$  zu  $N + M = 3 + 3 = 6$ ) auszuführen und die so erhaltenen Ergebnisse wieder zurückzuprojizieren. Dies kostet im Allgemeinen sehr viel Rechenzeit, da die Funktionen von der Ordnung  $N$  in die Ordnung  $N + M$  konvertiert werden müssen, dann folgt die Berechnung auf den zur Ordnung  $N + M$  gehörenden Knoten und schließlich noch die Rückprojektion von  $M + N$  nach  $N$ . Da sich zeigt, dass meist nur die höheren Ordnungen zu diesem Effekt beitragen, begnügt man sich jedoch mit einem deutlich kostengünstigeren Filter (siehe Abschnitt 3.1.2), der die höheren Funktionsanteile wegämpft.

### Die Wahl des Zeitschrittes

Die Wahl des geeigneten Zeitschrittes gestaltet sich bei den EULER-Gleichungen etwas schwieriger, da sich im allgemeinen Fall die maximale Geschwindigkeit, das heißt die Schallgeschwindigkeit, auch während der Simulation ändert. Des Weiteren beeinflussen Aliasfehler die Stabilität des Systems in nicht unbeträchtlichem Umfang, besonders wenn die Felder nur schwach räumlich aufgelöst sind. Es gibt Ansätze, bei denen dieser Zeitschritt dynamisch anhand der momentanen maximalen Schallgeschwindigkeit im System

angepasst wird:

$$\Delta t = \frac{1}{\max \lambda} \Delta t_{fix} \quad (3.3.6)$$

In dieser Arbeit wurde der Zeitschritt jedoch, soweit nicht anders genannt, festgehalten. Es zeigte sich, dass damit brauchbare Ergebnisse erzielt werden konnten. Nur wenn die Lösung schwach aufgelöst wurde, kam es zur Instabilität des Systems, die sich aber resistent gegen eine Verkürzung des Zeitschrittes zeigte.

### 3.3.1. Randbedingungen

Als Randbedingung kommen bei der EULER-Gleichung die „Slip“- und die „No-Slip“- Bedingungen zum Einsatz. Dabei bezeichnet erstere die klassische Reflexion der Teilchen am Rand. Darüber hinaus fordert letztere zusätzlich, dass die Tangentialkomponente der Geschwindigkeit identisch verschwindet. Die Berücksichtigung der Randbedingungen erfolgt hier über die Einführung virtueller äußerer Knoten.

**Tabelle 3.3.1.:** Randbedingungen für die EULER-Gleichung

	$\rho^+$	$U^+$	$\rho^+ \vec{v}^+$
SLIP	$\rho^-$	$U^-$	$\rho^- \vec{v}^- - 2(\vec{n}^- \cdot \rho^- \vec{v}^-) \cdot \vec{n}^-$
NOSLIP	$\rho^-$	$U^-$	$-\rho^- \vec{v}^-$

### 3.3.2. Konvergenztests

Auch bei den EULER-Gleichungen wurde zunächst ein Konvergenztest durchgeführt. Dazu wird hier der „Isentrope Wirbel“ behandelt [16, 71], da dies eines der wenigen Systeme ist, für die eine analytische Lösung bekannt ist. Hierbei handelt es sich um einen Wirbel, der sich in einem Geschwindigkeitsfeld mitbewegt. Die analytische Lösung beträgt (mit  $\gamma$  wieder als adiabatischem Exponenten,  $\beta$  als Wirbelstärke und  $T$  als Temperatur des Fluids<sup>32</sup>):

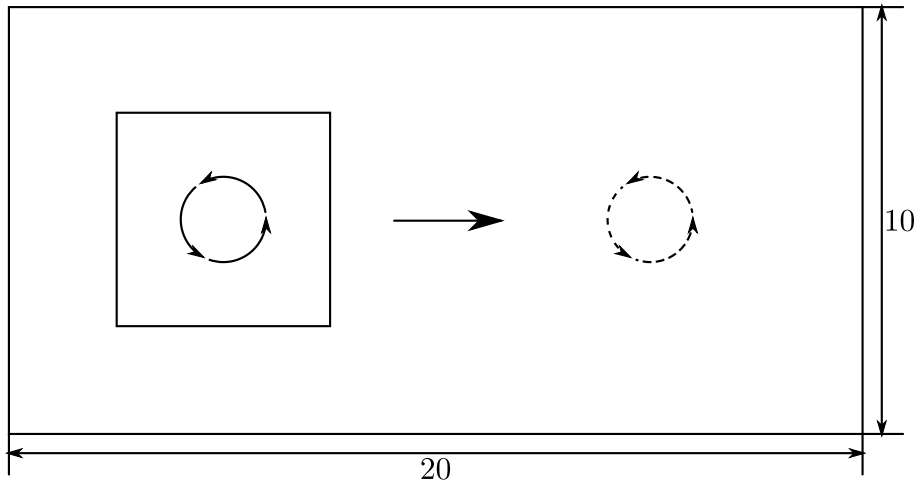
$$\rho = T^{1/(\gamma-1)} = (T_\infty + \delta T)^{1/(\gamma-1)} = \left( T_\infty - \frac{(\gamma-1)\beta^2}{16\gamma\pi^2} e^{2(1-\delta r^2)} \right)^{1/(\gamma-1)} \quad (3.3.7a)$$

$$v_x = v_{x,\infty} + \delta v_x = v_{x,\infty} - \frac{\beta}{2\pi} e^{1-\delta r^2} \delta y \quad (3.3.7b)$$

$$v_y = v_{y,\infty} + \delta v_y = v_{y,\infty} + \frac{\beta}{2\pi} e^{1-\delta r^2} \delta x \quad (3.3.7c)$$

$$p = \rho^\gamma \quad (3.3.7d)$$

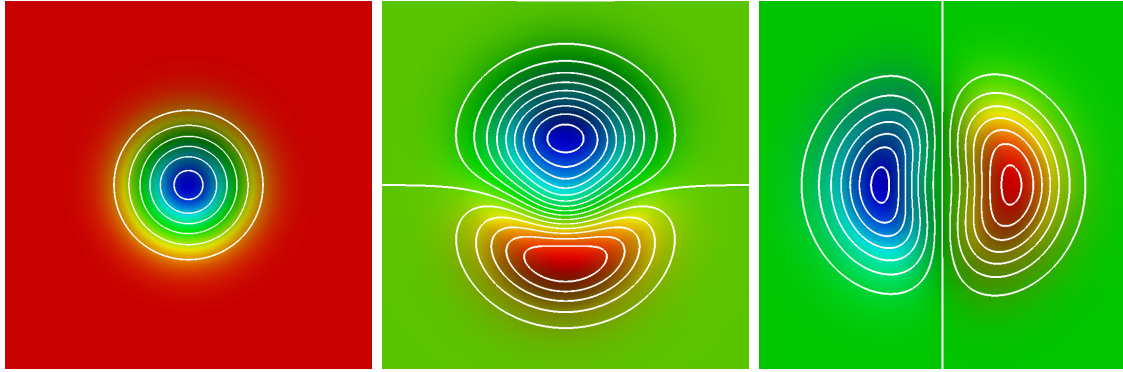
Des Weiteren ist  $\delta r^2 = \delta x^2 + \delta y^2$ ,  $\delta x = x - x_0$ ,  $\delta y = y - y_0$  und  $\vec{r}_0 = (x_0(t), y_0(t))^T$  die momentane Position des Wirbels. Die Werte für das Geschwindigkeitsfeld betragen für die Temperatur  $T_\infty = 1$  und die Geschwindigkeit  $\vec{v}_\infty = \vec{e}_x$ . Der Wirbel hatte eine Stärke von  $\beta = 5$ . Diese Lösung ist in Abbildung 3.3.3 für die Dichte  $\rho$  und die Stromdichte  $[\rho \vec{v}]$  skizziert. Das Simulationsgebiet ist in Abbildung 3.3.2 dargestellt. Für den Test startete



**Abbildung 3.3.2.:** Skizze des Simulationsgebietes des isentropen Wirbels. Gezeigt ist die Position des Wirbels zu Beginn und Ende der Simulation sowie der Rahmen, welcher die Funktionsauschnitte aus Abbildung 3.3.3 beinhaltet.

<sup>32</sup>Bei diesem System (isentropen Prozess und ideales Fluid) beträgt  $T = p/\rho$  und  $p/\rho^\gamma = 1$ .

der Wirbel in der linken Hälfte und es wurde bis  $\Delta t = 10$  gerechnet, so dass der Wirbel in der rechten Hälfte angekommen ist. Als Randbedingung wurde ein offenes Gebiet gewählt, was durch Vorgabe der exakten analytischen Lösung auf dem Rand  $\partial\Omega_h$  verwirklicht wurde. Eine graphische Darstellung der Dichte und der Stromdichte in der Startkonfiguration findet sich in Abbildung 3.3.3. Als Zeitschritt wurde hier  $c_{CFL} = 0.26c_{krit.}$  gewählt, was



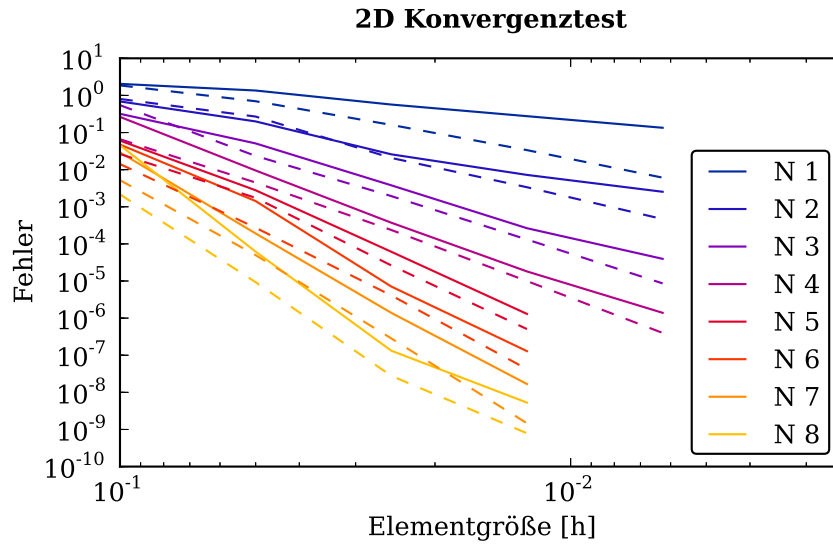
**Abbildung 3.3.3.:** Dichte,  $x$ - und  $y$ -Komponente der Stromdichte. Die Grunddichte beträgt 1, die Grundgeschwindigkeit  $\vec{v} = v_x = 1$ . Die Isokonturlinien haben einen Abstand von 0.1. Die Bilder sind 5x5 Einheiten groß.

zusammen mit einem Filter ( $s_{Filter} = 8$ ,  $\alpha_{Filter} = 0.5\%$ ) zu stabilen Ergebnissen geführt hat. Dabei ergibt sich die CFL-Konstante aus  $0.8/c_{max}$  mit  $c_{max} \approx 3.042$  der maximalen Schallgeschwindigkeit im System (siehe Gleichung 3.3.5 und 3.3.7)

Als Maß für den Fehler wurde die Abweichung in der Energie gewählt:

$$Fehler = \max_{N_p, N_e} |U^{analytisch} - U^{DG}| \quad (3.3.8)$$

Das Ergebnis ist in Abbildung 3.3.4 dargestellt. Als erstes fällt wieder auf, dass mit höherer Ordnung beziehungsweise feiner werdender Diskretisierung, der Fehler abnimmt. Des Weiteren nimmt er als Funktion der Elementgröße umso stärker ab, je höher die gewählte Ordnung ist. Dies ist wieder eine Folge der  $hp$ -Konvergenz (siehe auch Gleichung 3.2.9). Darüber hinaus sind die Systeme, bei welchen die „exakte“ Integration verwendet wurde, um bis zu 1,5 Größenordnungen genauer als die Systeme, bei denen nur mittels der Interpolationsknoten gerechnet wurde. Dieser Unterschied ist insbesondere bei niedrigen Ordnungen ( $N = 1$ ) beziehungsweise bei Systemen mit „schlechter“ Ortsauflösung (Elementgröße  $h > 5 \cdot 10^{-2}$ ) zu erkennen. Die Ursache ist in der schlechten räumlichen Überbestimmtheit der Interpolationsknoten zu sehen, die als Genauigkeitsreserve für die nichtlinearen Operationen während des Lösen der EULER-PDG benötigt wird. Im Allgemeinen überwiegt jedoch der Zeitvorteil der Interpolationsknoten gegenüber dem Genauigkeitsvorteil



**Abbildung 3.3.4.:** Konvergenzanalyse isentroper Vortex 2D. Die durchgezogenen Linien wurden unter Verwendung der Approximationsknoten und eines schwachen Filters bestimmt, die gestrichelten Linien mittels der im Text beschriebenen Projektionsoperatoren. Dabei kann man feststellen, dass die „exakte“ Integration, je nach Ordnung, um bis zu einer Größenordnung genauere Ergebnisse liefert.

der „exakten“ Integration.

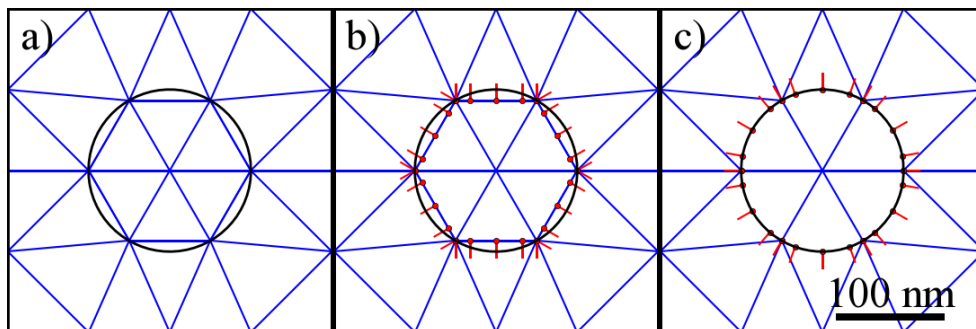


## 4. Gekrümmte Elemente

Nachdem im letzten Kapitel eine effiziente Variante der DG-Methode vorgestellt wurde, geht es in diesem Kapitel um eine weitere Verallgemeinerung, die krummlinigen Elemente. Bei jeder dieser Verallgemeinerungen, die in der Regel einen nicht unbeträchtlichen Mehraufwand mit sich bringen, gilt es immer abzuwägen, ob sich der Mehraufwand lohnt. Dass dies tatsächlich der Fall sein kann, soll durch einige Beispiele illustriert werden. Aber zuerst sollen die krummlinigen Elemente motiviert werden.

### Warum krummlinige Elemente?

Wenn man ein System mit krummlinigen Begrenzungsflächen, wie zum Beispiel Goldnanopartikel, mittels DG berechnen will, so muss es nach dem vorangegangenen Abschnitt in Teilgebiete zerlegt werden, wobei aus bereits erläuterten Gründen die Verwendung von d-Simplexen zweckmäßig ist. Schaut man sich Abbildung 4.0.1 an, so fällt auf, dass das Par-



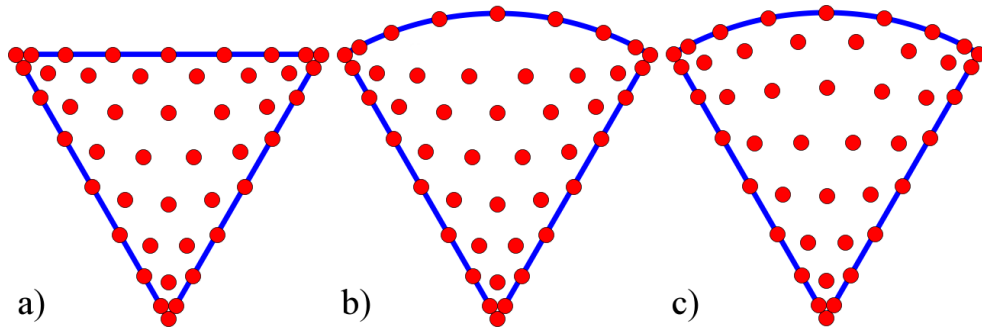
**Abbildung 4.0.1.:** Motivation runde Elemente. Linke Seite: Zerlegung eines Kreises in Dreiecke. Mitte: Die aus der Zerlegung resultierende Position der Randknoten inklusive der Richtung der Normalenvektoren. Rechte Seite: Korrigierte Position der Randknoten und Richtung der Normalenvektoren.

tikel nicht genau wiedergegeben wird. Dabei sind hier zwei Unstimmigkeiten festzuhalten: erstens, dass die Randknoten nicht auf dem zu beschreibenden Rand liegen, und zweitens, dass auch die Normalenvektoren, die zur Berechnung der Stetigkeitsbedingung (genauer der Oberflächenflussintegrale) notwendig sind, nicht in die richtige Richtung weisen.

Insbesondere existieren für jede Ecke/Kante, im Gegensatz zum Ausgangsobjekt, mehrere Normalenvektoren, was zwangsläufig zu Irritationen führt. Die Idee der krummlinigen Elemente ist nun, diese beiden Probleme zu korrigieren.

## 4.1. Allgemeines Verfahren

Um nun zu den krummlinigen Elementen zu gelangen, sind drei Schritte nötig. Erstens wird eine Darstellung der Knoten auf dem Rand benötigt, danach erfolgt die Aktualisierung der Volumenknoten und schließlich die Anpassung der Operatoren. In Abbildung



**Abbildung 4.1.1.:** Vorgehensweise zur Konstruktion der runden Elemente. Ausgehend vom ursprünglichen Element, wird als Erstes der Rand der Bogenlänge nach korrigiert, bevor als Zweites diese Verschiebung in das Innere des Elementes übertragen wird.

4.1.1 sind die ersten beiden Schritte schematisch dargestellt. Dabei ist darauf zu achten, dass die Abbildung von den linearen zu den krummlinigen Randknoten in der Weise geschieht, dass die Knotenabstände, bezogen auf die Bogenlänge, gleich bleiben, um die Interpolationseigenschaften nicht zu verschlechtern. Als nächstes werden die Änderungen analog zur Konstruktion der Volumenknoten wieder in das Innere hineingezogen (siehe Abschnitt 3.1.1). Die zugehörige Vektorverschiebungsfunktion  $\vec{w}_{e=1}^{2D,c}$  (siehe Gleichung 3.1.23) lässt sich mit Hilfe der alten Randknoten  $\vec{r}_{i,e}$ , den neuen Randknoten  $\vec{r}_{i,e}^c$  und den 1D-Interpolationspolynomen  $l_{i,e}^{1D}$  (am Beispiel der ersten Kante) wie folgt:

$$\vec{w}_{e=1}^{2D,c}(r, s) = \sum_i (\vec{r}_{i,e=1}^c - \vec{r}_{i,e=1}) l_{i,e=1}^{1D} (\lambda_3 - \lambda_2) \quad (4.1.1)$$

konstruieren. Der einzige Unterschied besteht darin, dass die Verschiebungsfunktion nun neben Komponenten entlang der Oberfläche auch solche senkrecht dazu beinhaltet.

Nachteilig bei den krummlinigen Elementen ist, dass die Transformation vom Referenzelement auf das physikalische Element nicht mehr so einfach vonstatten geht, da die

JACOBI-Matrix nun nicht mehr konstant für alle Knotenpunkte ist.

Bevor nun die Konstruktion der Operatoren präsentiert wird, sei noch eine kurze Bemerkung über numerische Integration eingeschoben. Diese lautet in allgemeinsten Form:

$$\int_{\Omega} f(\vec{r}) = \sum_i w_i f(\vec{r}_i) \quad \vec{r}_i \in \Omega \quad (4.1.2)$$

$$= \sum_i J_i^k w_i f(\vec{r}_i) \quad \vec{r}_i \in I \quad (4.1.3)$$

Dabei sind die  $\vec{r}_i$  geeignete Punkte aus dem zu integrierenden Gebiet und die  $w_i$  die zugehörigen Wichtungen. Hier wurde zusätzlich noch die Transformation des Integrals auf das Referenzelement durchgeführt. Damit ist ein Integral nichts anderes als die Summe über die Interpolation der zu integrierenden Funktion. Die Wichtungen sind dann der Beitrag einer Interpolationsfunktion zum Gesamtintegral. Nicht von ungefähr entspricht auch die erste Zeile der VANDERMONDE-Matrix den Wichtungen<sup>1</sup>. Die Frage nach der Integrierbarkeit einer Funktion ist damit äquivalent zu deren Interpolierbarkeit. Die Stützpunkte sind demnach so zu bestimmen, dass die Funktion mit möglichst wenigen Punkten numerisch stabil interpoliert werden kann<sup>2</sup>.

Mit dieser Vorbetrachtung können nun die Operatoren konstruiert werden. Für die Massenmatrix folgt:

$$M_{ij}^k = \int_I l_i l_j J^k \quad (4.1.4)$$

$$= \sum_l l_i(r_l) l_j(r_l) w_l J^k(r_l) \quad (4.1.5)$$

und unter Verwendung der VANDERMONDE-Matrix  $V_{lj}^c = l_j(r_l)$  gebildet an den Integralknoten:

$$M^k = (V^c)^T \text{diag}(w J^k) V^c \quad (4.1.6)$$

<sup>1</sup>Die erste Zeile enthält den Anteil des konstanten JACOBI-Polynoms.

<sup>2</sup>Passende Integrationsregeln finden sich in [13, 74]

Die Berechnung der Ableitungsmatrizen  $\vec{D}^k$  erfolgt sinngemäß:

$$S_{x,ij}^k = \int_{\Omega^k} \partial_x l_i l_j \quad (4.1.7)$$

$$= \int_{\Omega^k} (\partial_x r \partial_r + \partial_x s \partial_s + \partial_x t \partial_t) l_i l_j \quad (4.1.8)$$

$$= \int_I J^k (\partial_x r \partial_r + \partial_x s \partial_s + \partial_x t \partial_t) l_i l_j \quad (4.1.9)$$

$$= \sum_l w_l J^k (\partial_x r \partial_r + \partial_x s \partial_s + \partial_x t \partial_t) l_i(r_l) l_j(r_l) \quad (4.1.10)$$

Das lässt sich wieder auf Matrixschreibweise bringen:

$$S_x^k = (V^c)^T \text{diag}(w J^k) (\text{diag}(\partial_x r) D_r^c + \text{diag}(\partial_x s) D_s^c + \text{diag}(\partial_x t) D_t^c) \quad (4.1.11)$$

Dabei sind die  $D_I^c = V_I^c V^{-1}$  mit  $V_{I,ij}^c = \nabla_I \psi_j(r_i^c)$  wie bei den linearen Elementen konstruiert, nur jetzt eben an den neuen Integrationspunkten  $r^c$ . Die fehlenden Matrizen für  $S_y^k$  und  $S_z^k$  erhält man entsprechend. Die Verknüpfung zur Ableitungsmatrix  $\vec{D}^{k,c}$  erfolgt über die inverse Massenmatrix:

$$\vec{D}^{k,c} = (M^{k,c})^{-1} \vec{S}^{k,c} \quad (4.1.12)$$

Für die Integration an der Oberfläche geht man ähnlich vor. Zuerst werden die Elementknoten (und zwar alle, nicht nur die auf dem Rand) in die Integrationsknoten mittels der Interpolationsmatrix  $V_{ij}^g = l_j(r_i^g)$  überführt. Danach erfolgt die Flussberechnung, wie in den entsprechenden Abschnitten 3.2 und 3.3 erwähnt, und schließlich die Abintegration mittels:

$$F^g = (V^g)^T \text{diag}(w^g J_g^k) \quad (4.1.13)$$

Hierbei ist zu beachten, dass der Oberflächenintegraloperator nun mit  $(M^{k,c})^{-1}$  zu multiplizieren ist und nicht mehr mit  $(M^k)^{-1}$ .

### Räumliche Ableitung in der schwachen Form

Während die Massenmatrizen (für das Volumen und die Oberflächen) für die starke und die schwache Form der PDGL-Systeme identisch sind, muss bei der schwachen Form der räumlichen Ableitung die transponierte Darstellung verwendet werden. Damit lautet die Ableitung in schwacher Form:

$$D_{x,w} = (M^{k,c})^{-1} (D_r^T \text{diag}(w^c \partial_x r) + D_s^T \text{diag}(w^c \partial_x s) + D_t^T \text{diag}(w^c \partial_x t)) \quad (4.1.14)$$

Die anderen Richtungsableitungen ergeben sich wieder sinngemäß.

### Bemerkung zu Rechenzeit

Ähnlich zu Abschnitt 3.1.3 soll auch hier eine Abschätzung der Rechenzeit  $t_{CPU}$  erfolgen. Dazu wird Formel 3.1.47, wie folgt, erweitert:

$$t_{CPU} = t_{\Omega_h^k} \tilde{N}_t (N_e + c_{N_{ce}} N_{ce} + c_+) \quad (4.1.15)$$

Dabei ist  $N_{ce}$  die Anzahl der gekrümmten Elemente und  $c_{N_{ce}}$  der zusätzliche Aufwand für ein rundes Element, bezogen auf den Aufwand für ein lineares Element<sup>3</sup> und  $\tilde{N}_t = c_{N_t} N_t$  berücksichtigt eine Änderung der Zeitschrittzahl als Folge der Anpassung des kritischen Zeitschrittes  $\Delta t$ . Das heißt, die runden Elemente können auf zwei Wegen zu einer Vergrößerung der Rechenzeit beitragen. Einerseits, da durch die runden Elemente mehr Zeitschritte notwendig sein können, als für lineare Elemente, und andererseits, da die Behandlung eines runden Elementes aufwändiger ist, als die Behandlung eines linearen Elementes. Auf der anderen Seite könne runde Elemente insgesamt zu einer Einsparung der Rechenzeit führen, wenn dadurch mit einer gröberen Diskretisierung gearbeitet werden kann. Diese gröbere Diskretisierung hat die Vorteile, dass der kleinste Zeitschritt größer gewählt werden kann (aufgrund des größeren kleinsten Elementes), was unter Umständen die Anpassung von  $c_{N_t}$  kompensiert und das weniger Elemente behandelt werden müssen, so dass auch hier der Mehraufwand  $c_{N_{ce}}$  ausgeglichen werden kann.

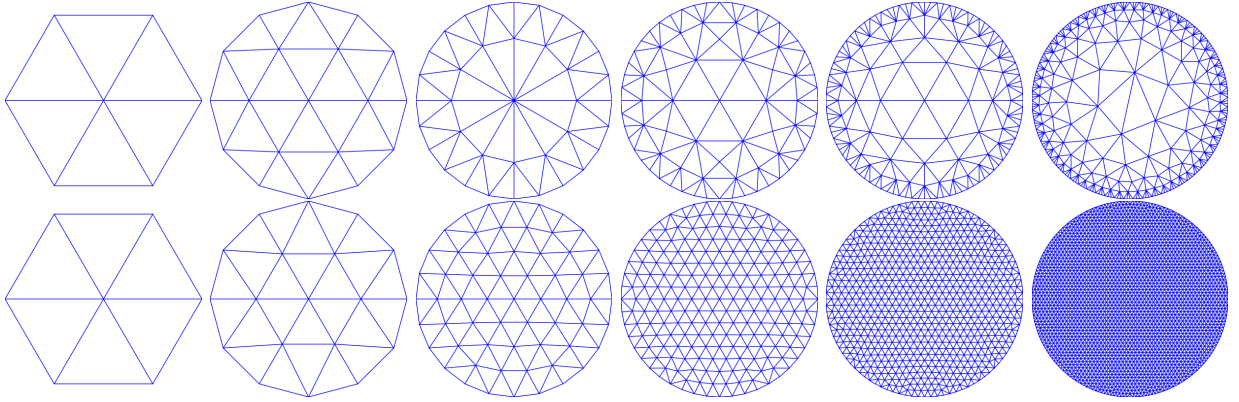
---

<sup>3</sup>Eine alternative Darstellung behandelt die linearen und die runden Elemente unabhängig mit:

$$t_{CPU} = t_{\Omega_h^k} \tilde{N}_t (N_{le} + \tilde{c}_{N_{ce}} N_{ce} + c_+)$$

Dabei ist nun  $N_{le}$  die Anzahl der linearen Elemente mit  $N_{le} = N_e - N_{ce}$  und  $\tilde{c}_{N_{ce}} = c_{N_{ce}} + 1$ .

## 4.2. Gekrümmte Elemente in 2D



**Abbildung 4.2.1.:** Netze 2D-Kreiskavität. Die Auflösung des Randes wurde von sechs Randelementen auf 192 Randelemente vergrößert. Dabei stellt die obere Reihe eine randselektive Verfeinerung des Netzes und die untere Reihe eine uniforme Verfeinerung des Netzes dar.

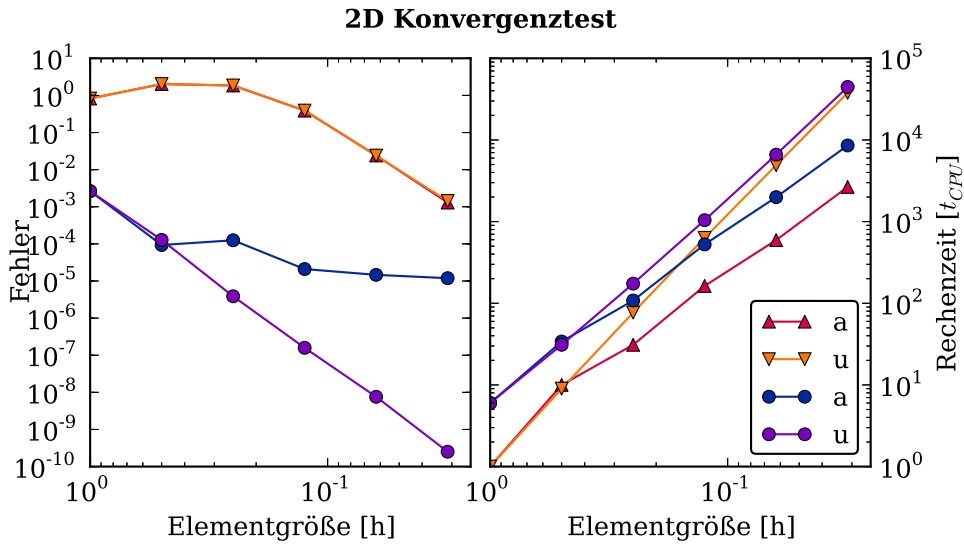
Für den 2D benötigt man für eine möglichst genaue und stabile Darstellung der Oberfläche nur eine Parameterdarstellung des Randes als Funktion der Bogenlänge um die Oberflächenpunkte zu bestimmen. Die Abwicklung der Kurve auf das Referenzelement (1D - Linie) ist dann einfach. Für die in der Arbeit verwendeten Netze wurden nur Kreisausschnitte verwendet. Die zugehörigen Knotenpunkte wurden vom hauseigenen Netzgenerator A.1 analog zu Abschnitt 4.1 bestimmt<sup>4</sup>.

Als erster Test für die Implementation diente eine kreisförmige Kavität. Dazu wurden je sechs Netze mit feiner werdendem Rand, einerseits mit uniformer Elementgröße und andererseits mit selektiver Elementgröße betrachtet (siehe Abbildung 4.2.1). Als Ausgangsfeld diente die Grundmode der Kavität (eine „H“-polarisierte zylindrische Besselfunktion nach Abschnitt 2.1.4). Nach 100 Periodendauern wurde der maximale Fehler der  $E_z$ -Komponente durch Vergleich mit der analytischen Lösung bestimmt:

$$Fehler = \max_{Ne, Np} |E_z^{analytisch} - E_z^{DG}| \quad (4.2.1)$$

Als Basisfunktion wurde die Ordnung  $N = 4$ , einmal mit gekrümmten sowie einmal mit linearen Elementen verwendet. Das Ergebnis ist in Abbildung 4.2.2 dargestellt. Als erstes fällt auf, dass man mit runden Elementen eine optimale Konvergenz erreicht. Des Weiteren

<sup>4</sup>In [16] ist die Projektion von Rand- zu Volumenknoten nur in 2. Ordnung exakt, in höheren Ordnungen verzieht die Projektion, da nicht mit den baryzentrischen Koordinaten gearbeitet wird. Für kleine Änderungen zur quadratischen Näherung (z.B. alle Netze in dieser Arbeit) war der Unterschied jedoch vernachlässigbar.



**Abbildung 4.2.2.:** Konvergenzanalyse der 2D-Kreiskavität. Linke Seite: Fehler, rechte Seite: Rechenzeit. Kreise: runde Elemente, Vierecke: lineare Elemente. Label „a“: adaptive Verfeinerung, Label „u“: uniforme Verfeinerung.

ren sieht man, dass für lineare Elemente die Volumendiskretisierung keine Rolle spielt, das heißt, die Fehlerkurven liegen genau aufeinander. Dies bedeutet, dass der ganze Fehler ausschließlich vom Rand herrührt. Zudem ist das feinste Netz mit linearen Elementen immer noch schlechter als das gröbste Netz mit runden Elementen, obwohl es die 32fache Anzahl an Randelementen zeigt, sodass schon jetzt die bessere Effizienz der runden Elemente deutlich wird. Differenzierter sieht es beim Vergleich zwischen uniformer und nichtuniformer Verfeinerung bei den runden Elementen aus. Hier wird recht schnell der Volumenanteil am Fehler durch die größeren Elemente im Inneren der selektiven Verfeinerung sichtbar, während die uniforme Verfeinerung eine Genauigkeit bis hin zur finiten Zahlenrepräsentation hinab ermöglicht.

### Bemerkung zur Rechenzeit

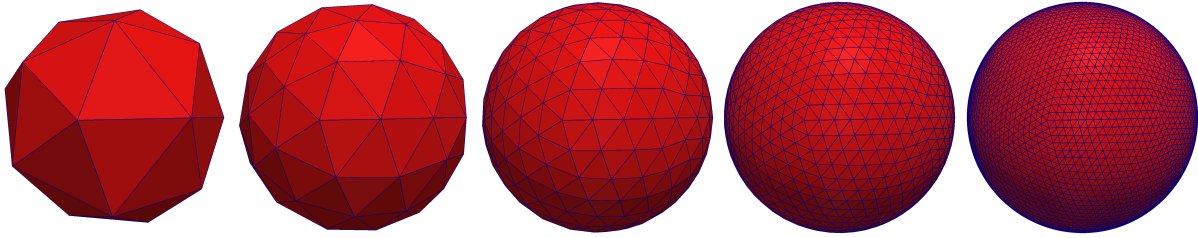
Für alle oben genannten Rechnungen war  $c_{Nt} \approx 1$ , da es sich bei den hier verwendeten runden Elementen um konvexe Verformungen der Elemente handelte. Der zusätzliche Mehraufwand lag bei  $c_{Nce} \approx 5$ . Damit ist ein rundes Element in der in dieser Arbeit verwendeten Implementation ca. sechsmal teurer wie ein lineares Element. Die normierten Rechenzeiten sind in Abbildung 4.2.2 zusammengetragen. Dabei fällt sowohl die niedrigere Rechenzeit für die adaptiv verfeinerten Netze im Vergleich zu den uniform verfeinerten Netzen auf, wie auch der zunehmende Angleich der Rechenzeiten von linearen und run-

den Elementen bei uniformer Verfeinerung der Netze. Die Ursache für letzteres liegt in der unterschiedlichen Zunahme der inneren und äußeren Elemente. So steigt die Anzahl der inneren Elemente mit  $(1/h)^2$ , während die runden Elemente nur mit  $1/h$  zunehmen, so dass sich mit fortschreitender Verfeinerung das Verhältnis von runden zu linearen Elementen zugunsten der linearen entwickelt, und so die Zeit, welche mit der Berechnung der runden Elemente verbracht wird, eine geringere Rolle spielt. Anders verhält es sich mit der selektiven Verfeinerung, bei der sowohl die Anzahl der Randelemente, wie auch die Anzahl der inneren Elemente im Wesentlichen linear mit  $1/h$  zunimmt.

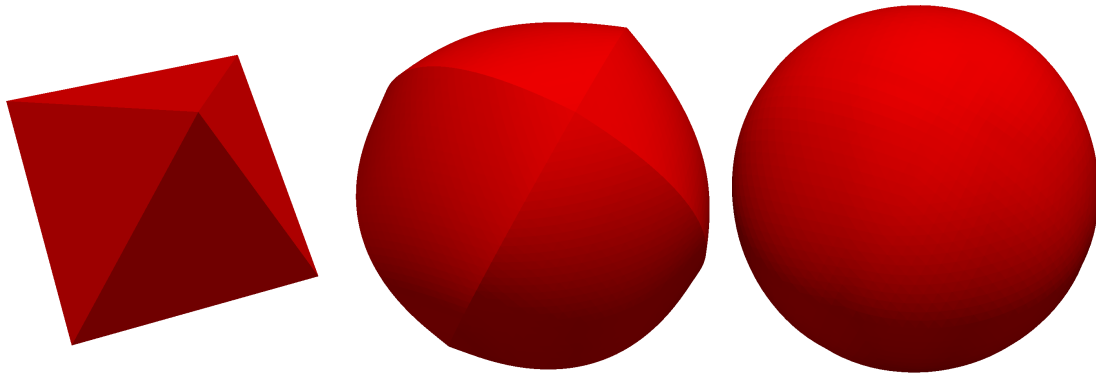
Vergleicht man nun noch das feinste adaptive Netz mit linearen Elementen mit dem größten uniformen Netz mit runden Elementen, da diese ungefähr den gleichen Fehler besitzen, so sind die runden Elemente bei gleichem Fehler ca. 443 mal schneller als die linearen Elemente.



### 4.3. Gekrümmte Elemente in 3D



**Abbildung 4.3.1.:** Netze der 3D-Kugelhöhle. Die Auflösung des Randes wurde von 42 Randelementen auf 20832 Randelemente verfeinert. Dabei stieg die Gesamtzahl der Elemente von 42 auf 172032 Elemente.

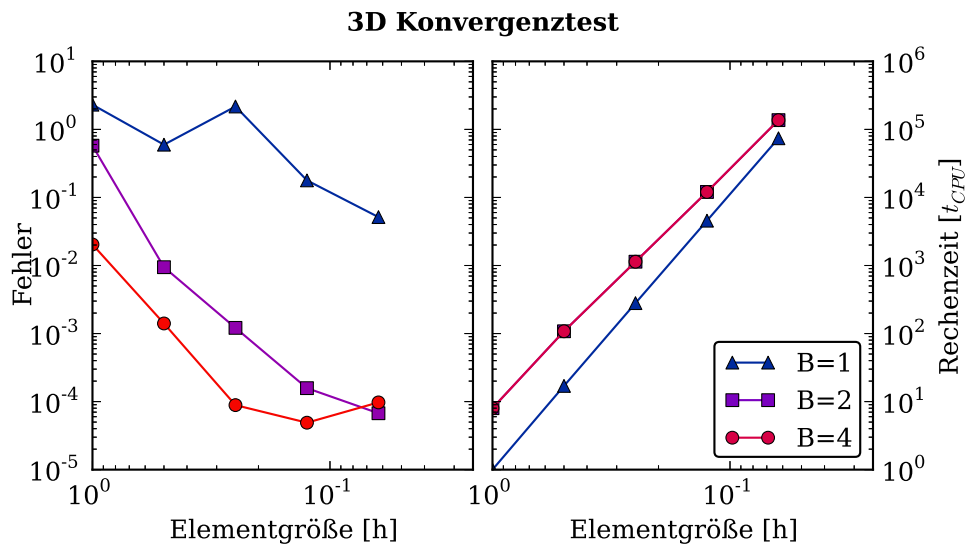


**Abbildung 4.3.2.:** Beispiel verschiedener Approximationsordnungen (von links nach rechts:  $B = 1, 2$  &  $4$ ) der Oberfläche einer Kugel am Beispiel der Umkugel eines Oktaeders. Die Güte (Ordnung  $B$ ) der Oberflächenapproximation darf nicht verwechselt werden mit der Ordnung  $N$  der Approximationspolynome für das elektromagnetische Feld.

Im Gegensatz zum Zweidimensionalen gestaltet sich die Gewinnung der Oberflächenkrümmung hier etwas schwieriger. Einerseits lieferte der verwendete Netzgenerator (NETGEN siehe [60]) nur eine Oberflächenrepräsentation bis zur zweiten Ordnung, das heißt, die Eckpunkte sowie die Kantenhalbierenden zwischen den Eckvertices. Andererseits existiert keine einfache Abwicklung einer allgemein gekrümmten Oberfläche auf die Ebene, so dass hier zwangsläufig Abstriche hinsichtlich der Qualität der Integration in Kauf genommen werden müssen. Darüber hinaus existieren im Dreidimensionalen tendenziell mehr Randelemente, da nun nicht nur Elemente als Rand gezählt werden welche über eine Randfläche verfügen, sondern auch jene welche über eine Randkante verfügen. Das heißt, im 3D zählen alle Elemente als runde Elemente, die über wenigstens eine Kante verfügen, die auf einer gekrümmten Oberfläche liegt<sup>5</sup>.

<sup>5</sup>Damit sind automatisch auch alle Elemente erfasst, bei denen eine Fläche, und damit mindestens drei

Als Testsystem dient nun eine kugelförmige Kavität. Es diene wieder die Grundmode als Anregung (eine „H“-polarisierte sphärische Besselfunktion nach Abschnitt 2.1.4). Um den Einfluss der Oberflächenrepräsentation auf das Ergebnis zu studieren, wurden in diesem Beispiel fünf Netze (siehe Abbildung 4.3.1) mit verschiedenen polynomialen Approximationsordnungen  $B = 1, 2, 4$  der Kugeloberfläche (siehe Abbildung 4.3.2) verwendet. Dabei kam zur Gewinnung der Oberflächenapproximation der in Abschnitt A.2 beschriebene Algorithmus zum Einsatz. Die Ordnung der Basisfunktionen war wieder  $N = 4$ . Der maximale Fehler der  $E_z$ -Komponente nach 100 Periodendauern ist in Abbildung 4.3.3 dargestellt. Es



**Abbildung 4.3.3.:** Konvergenzanalyse der 3D-Kugelkavität. Linke Seite: Fehler, rechte Seite: zugehörige Rechenzeit. Bei der Berechnung ist es für die runden Elemente unerheblich in welcher Genauigkeit die Oberfläche vorliegt. Daher ist die Rechenzeit für  $B = 2$  und  $B = 4$  identisch.

wird auch hier die Bedeutung der Oberflächenapproximation deutlich. So erzielt bei ansonsten gleichen Parametern die jeweils höhere Oberflächenapproximationsordnung (nicht zu verwechseln mit der Ordnung der Basispolynome) einen kleineren Fehler als die jeweils niedrigere. Des Weiteren fällt hier das Fehlerminimum bei  $10^{-4}$  auf, dessen Ursache nicht vollständig geklärt ist, da sich mit den linearen Elementen, aufgrund der Größe der Projekte und der zur Verfügung stehenden Rechenkapazität, keine derartig genauen Ergebnisse erzielen ließen. Auch war es mit höherer Ordnung nicht möglich ein genaueres Ergebnis zu erzielen, so dass die  $10^{-4}$  als unteres systembedingtes Limit gesehen werden kann. Des Weiteren ist auch hier das feinste Netz mit linearen Elementen deutlich schlechter, als das gröbste Netz mit runden Elementen und genauester Oberflächenrepräsentation.

Kanten, auf einer gekrümmten Oberfläche liegen.

**Bemerkung zur Rechenzeit**

Auch hier lag die Zeitkonstante  $c_{Nt} \approx 1$ , da auch in diesem Fall wieder nur konvexe Verformungen der Elemente stattfanden. Darüber hinaus ist es für die runden Elemente nicht von Bedeutung, welche Genauigkeit die zugrundeliegende Repräsentation des Randes besitzt, so dass die Rechenzeiten für  $B = 2$  und  $B = 4$  identisch sind. Des Weiteren lag der empirisch ermittelte Wert für  $c_{Nce}$  bei  $\approx 7$ , so dass ein rundes Element achtmal länger in der Berechnung dauert wie ein lineares Element. Die Ergebnisse der Rechenzeit in Abhängigkeit vom Netz sind in Abbildung 4.3.3 dargestellt. Auch hier wird wieder der Anstieg der Rechenzeit mit feiner werdendem Netz und der Angleich der Rechenzeiten zwischen den linearen und den runden Elementen deutlich. Ermittelt man nun den Geschwindigkeitsvorteil der runden gegenüber den linearen Elementen bei etwa gleichem Fehler, so erhält man für  $B = 4$  eine circa 9300 mal kürzere Rechenzeit und für  $B = 2$  immerhin noch eine 690 mal kürzere Rechenzeit.



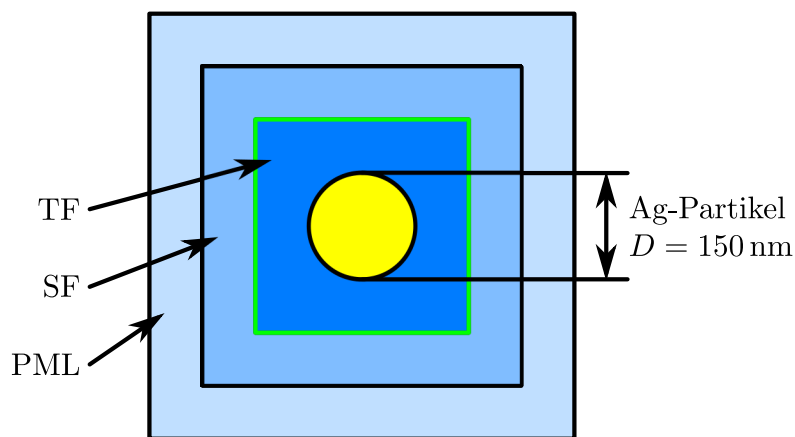
## 5. Ergebnisse

### 5.1. Plasmonische Nanostrukturen im 2D

Nachdem in den vorangegangenen Kapiteln die Grundlagen der DG-Methode erläutert wurden, sollen nun die ersten Anwendungen auf nanooptische Systeme folgen. Dabei wird im Folgenden der Auswirkung der runden Elemente in plasmonischen Systemen die meiste Aufmerksamkeit gewidmet.

#### 5.1.1. Partikel

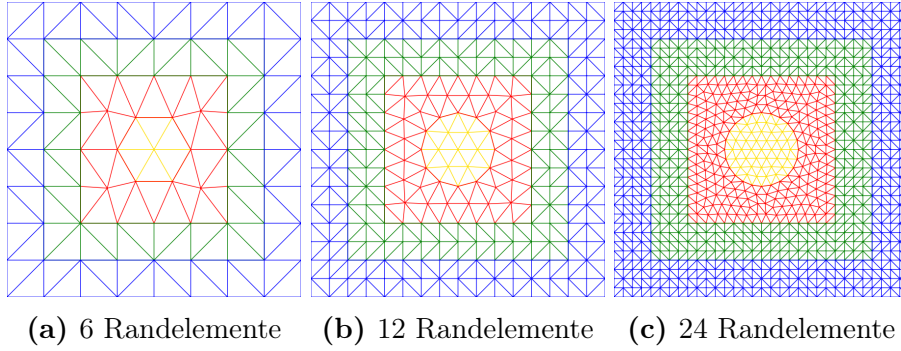
Als erstes folgt die Betrachtung eines unendlich langen zylindrischen Stabes (2D-Kugel). Es ist eines der einfachsten plasmonischen Systeme mit bekannter analytischer Lösung ([4, 42]). Das hier behandelte System ist in Abbildung 5.1.1 skizziert. Der Durchmesser der



**Abbildung 5.1.1.:** Skizze der 2D-Kugel. Das System besteht aus insgesamt vier Bereichen: der 2D-Kugel, dem TF-Bereich, dem SF-Bereich und dem abschließenden PML

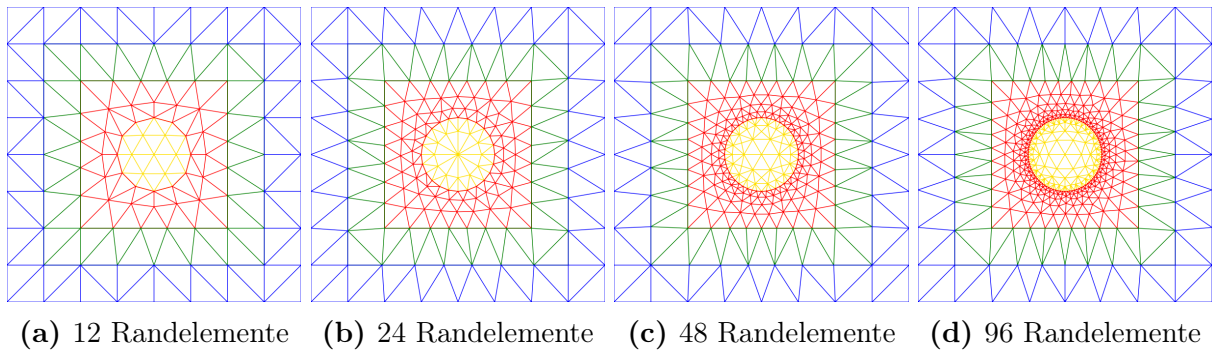
„2D-Silberkugel“ beträgt 150 nm, das Material entspricht dem in Tabelle 2.3.1 aufgeführten DRUDE-Silber. Da mit den Erkenntnissen aus den vorangegangenen Kapiteln die Oberfläche der Kugel den kritischen Teil der Simulation bildet, werden hier zwei verschiedene Diskretisierungsvarianten betrachtet. Einerseits drei uniforme Verfeinerungen des Netzes

mit sechs bis 24 Randelementen (siehe Abbildung 5.1.2), andererseits fünf oberflächenselektive Verfeinerungen mit 6 bis 96 Randelementen (siehe Abbildung 5.1.3). Es wurden

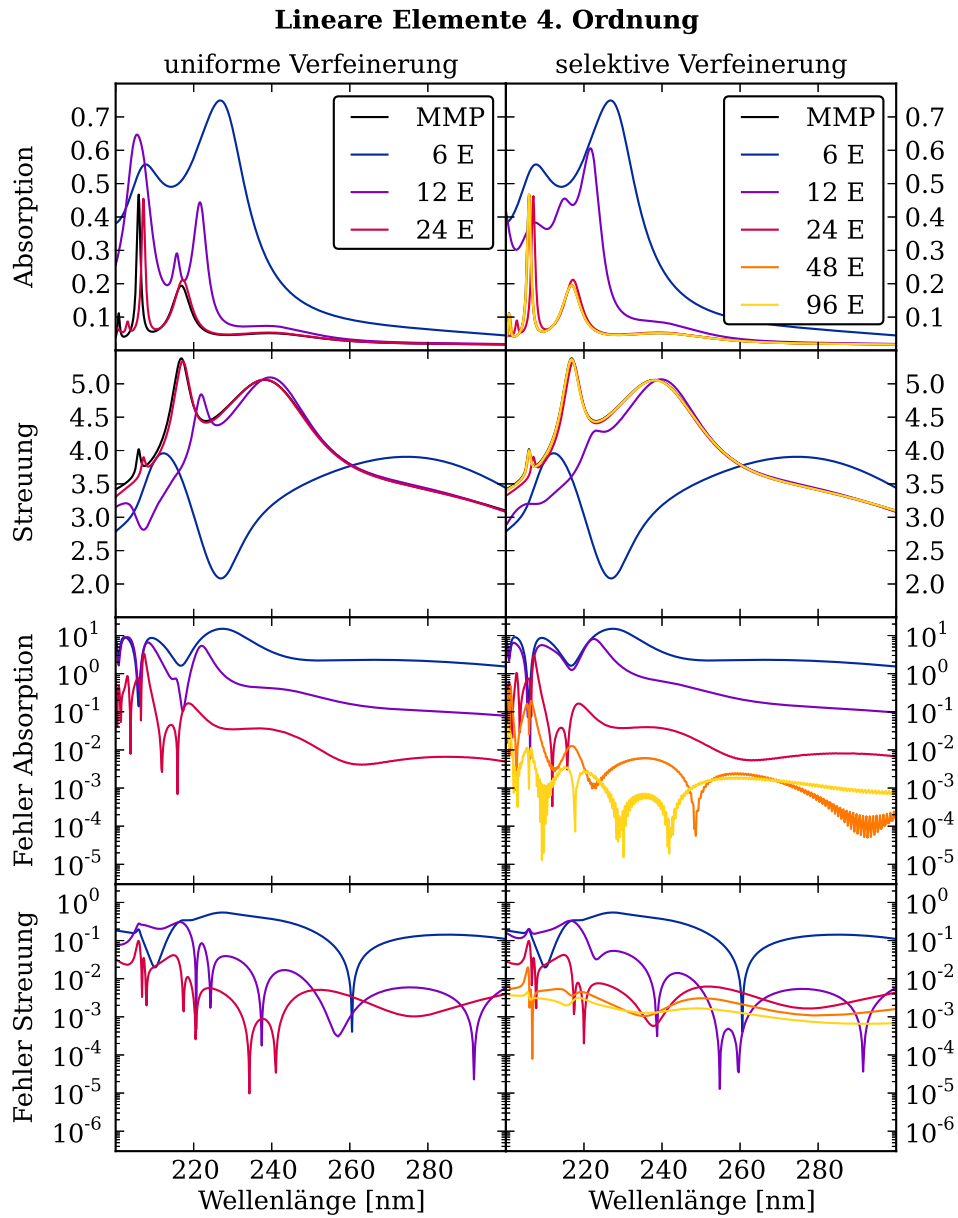


**Abbildung 5.1.2.:** Serie uniform verfeinerter Netze. Die Farben grenzen die verschiedenen Simulationsgebiete (Gelb: Partikel, Rot: TF, Grün: SF und Blau: PML-Bereich) voneinander ab.

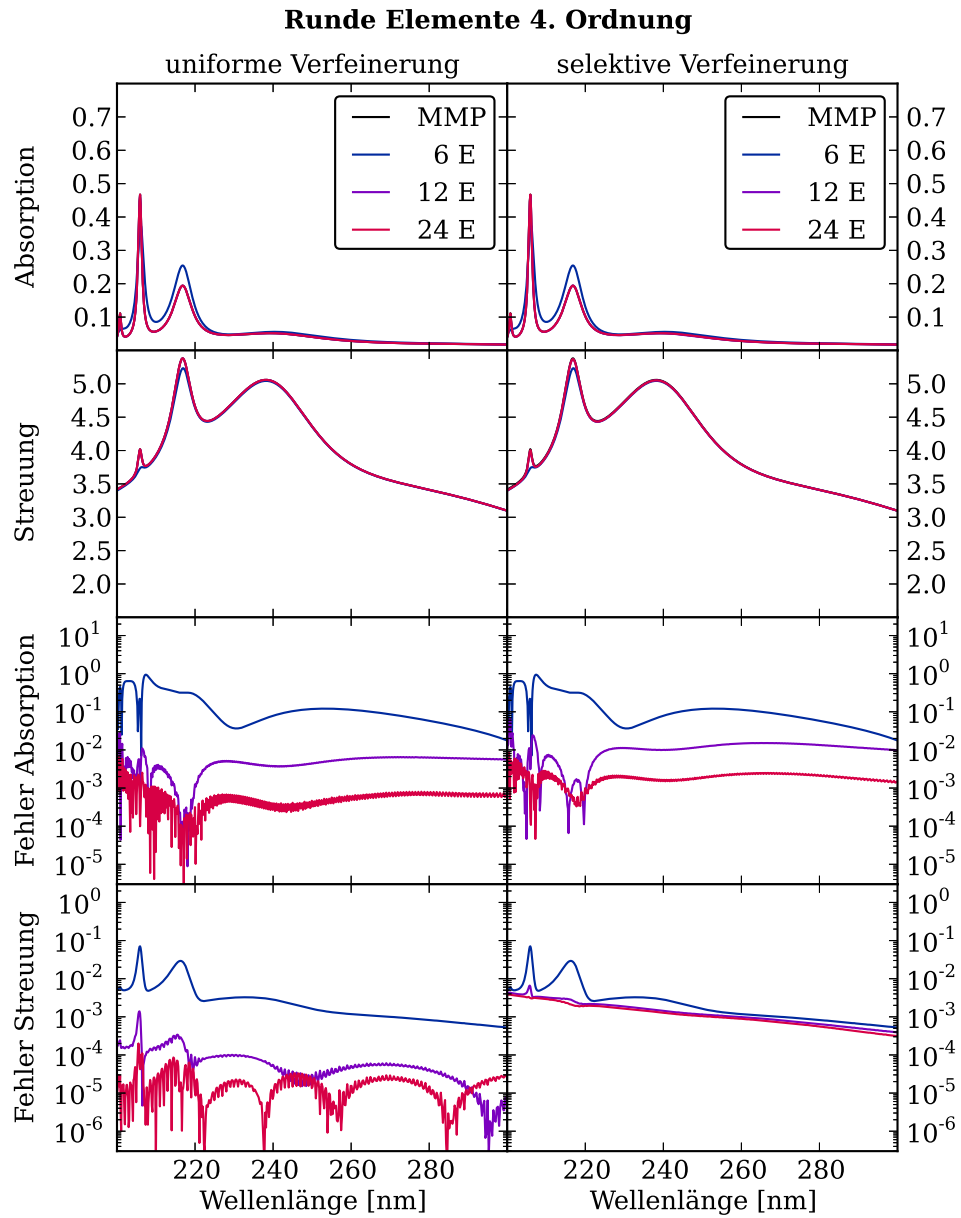
bei diesen Rechnungen Polynome 4. bzw. 8. Ordnung mit linearen und runden Elementen verwendet. Als Anregung diente ein  $TE$ -polarisierter differenzierter GAUSSscher Puls (siehe Abschnitt 2.6.3 und Gleichung 2.6.4) mit einer Mittenwellenlänge von 240 nm und einer Propagation in  $+x$ -Richtung. Während der Simulation wurde das Spektrum an 500 Punkten im Bereich von 200 nm bis 300 nm Wellenlänge aufgezeichnet. Als Referenzmethode wurde hier MMP (siehe [14]) verwendet<sup>1</sup>.



**Abbildung 5.1.3.:** Serie selektiv verfeinerter Netze. Das Netz mit 6 Elementen wurde weggelassen, da es identisch mit Abbildung 5.1.2a ist. Die Farben grenzen die verschiedenen Simulationsgebiete (Gelb: Partikel, Rot: TF, Grün: SF und Blau: PML-Bereich) voneinander ab.



**Abbildung 5.1.4.:** Absorptions- und Streuspektren sowie deren relative Fehler für die 2D-Kugel, gerechnet mit linearen Elementen, für verschiedene Elementzahlen und Netz-Verfeinerungsstrategien.



**Abbildung 5.1.5.:** Absorptions- und Streuspektren sowie deren relative Fehler für die 2D-Kugel, gerechnet mit runden Elementen, für verschiedene Elementzahlen und Netz-Verfeinerungsstrategien.



### Feldüberhöhung 4. Ordnung

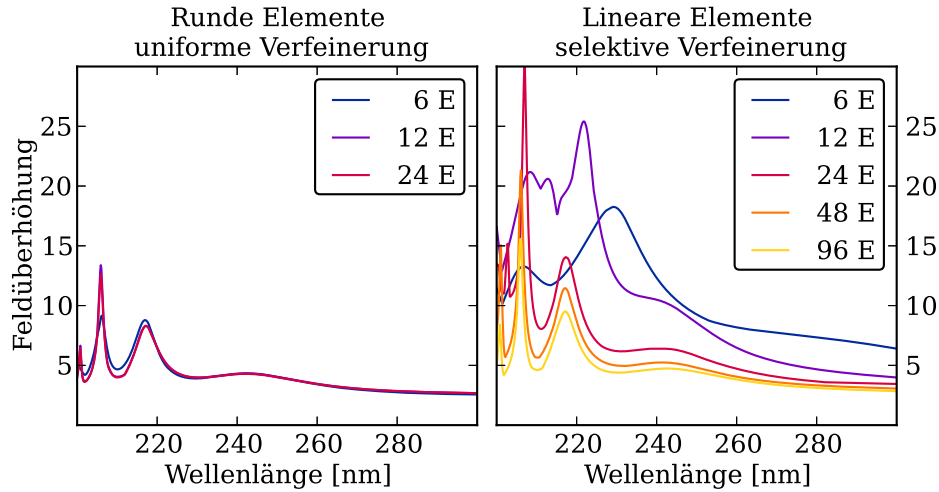


Abbildung 5.1.6.: Feldüberhöhung für runde und lineare Elemente

### Uniforme vs. nichtuniforme Verfeinerung

Zuerst werden die beiden Verfeinerungsvarianten (uniform und selektiv) für die linearen Elemente betrachtet. Das Ergebnis findet sich in Abbildung 5.1.4. Zuerst fällt auf, dass beide Varianten mit kleiner werdenden Oberflächenelementen gegen die Referenzlösung konvergieren und ab ca. 24 Oberflächenelementen ein brauchbares Ergebnis erzielen. Nichtsdestotrotz zeigt sich für die Rechnungen mit 12 Oberflächenelementen, dass beide Lösungen, obwohl sie dasselbe geometrische System (ein Zwölfeck) beschreiben, nicht übereinstimmen. Ursache dafür ist die Feldsingularität an den Ecken (ähnlich zu den 6 beziehungsweise 24 Ecken in Abbildung 5.1.8). Dies führt zu einer Art GIBBS-Phänomen<sup>2</sup>, welches das Spektrum beeinflusst. Weiterhin beobachtet man im Fehlerbild, dass die Absorption mit zunehmender Anzahl der Oberflächenelemente immer genauer wird. Anders verhält es sich bei der Streuung, welche bei ca.  $10^{-3}$  (bis auf einzelne Ausnahmen) ihren niedrigsten Fehler besitzt. Dass dieser nicht weiter sinkt, hängt nun hauptsächlich mit der Volumendiskretisierung zusammen, das heißt, die vom Partikel abgestrahlten Wellen können nicht genauer aufgelöst werden.

Als nächstes folgen analoge Betrachtungen zu den Rechnungen mit den runden Elementen (siehe Abbildung 5.1.5). Auch hier sieht man wieder eine mit feiner werdendem Netz genauer werdende Lösung. Am beeindruckendsten ist jedoch, dass selbst mit einer

<sup>1</sup>Prinzipiell kann man dafür auch die analytische MIE-Streuung ([4, 42]) verwenden. Einfachheitshalber wurde hier auf MMP zurückgegriffen. Die mit MMP erzielten Ergebnisse sind auf mindestens sechs signifikante Digits konvergiert.

<sup>2</sup>Analog zur FOURIER-Reihe von unstetigen Funktionen, kann es in der Umgebung singulärer Punkte zu einem Überspringen der Funktion kommen.

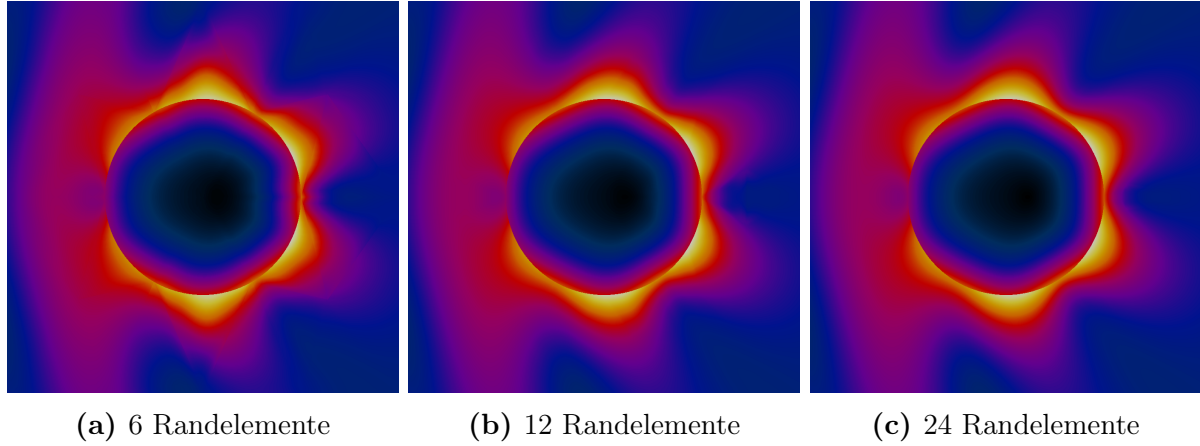
„sechseckigen“ Kugel sich bereits brauchbare Resultate erzielen lassen. Dabei fällt auf, dass die Ergebnisse für die Streuung sogar noch genauer ausfallen als für die Absorption. Die Erklärung folgt aus den verschiedenen notwendigen Ortsauflösungen für die beteiligten Wellenlängen des elektromagnetischen Feldes. So ist für die Propagation des Feldes, und damit für die Berechnung der Streuung, im Wesentlichen die Wellenlänge des Lichtes im Umgebungsmedium maßgebend, während für die Absorption die Propagation des Feldes auf der Oberfläche, und damit die Plasmonwellenlänge, die kritische Dimension darstellt. Diese beträgt nach Gleichung 2.4.2 nur ca.  $\frac{2}{3}$  der Luftwellenlänge. So sieht man, dass mit besserer Oberflächendiskretisierung der Fehler in der Absorption für beide Netztypen etwa gleich stark sinkt. Demgegenüber bleibt der Fehler für die Streuung für die selektiven Netze ungefähr konstant und verringert sich nur für die uniforme Verfeinerung. Nichtsdestotrotz zeigt die selektive Verfeinerung das ausgewogenere Fehlerbild, da hier Absorptions- und Streufehler ungefähr gleich groß sind. Darüber hinaus ist bei dem Streufehler sehr gut der Zusammenhang zwischen Elementgröße und Wellenlänge zu sehen. So steigt der Fehler für kürzere Wellenlängen beziehungsweise sinkt für größere.

### Runde vs. lineare Elemente

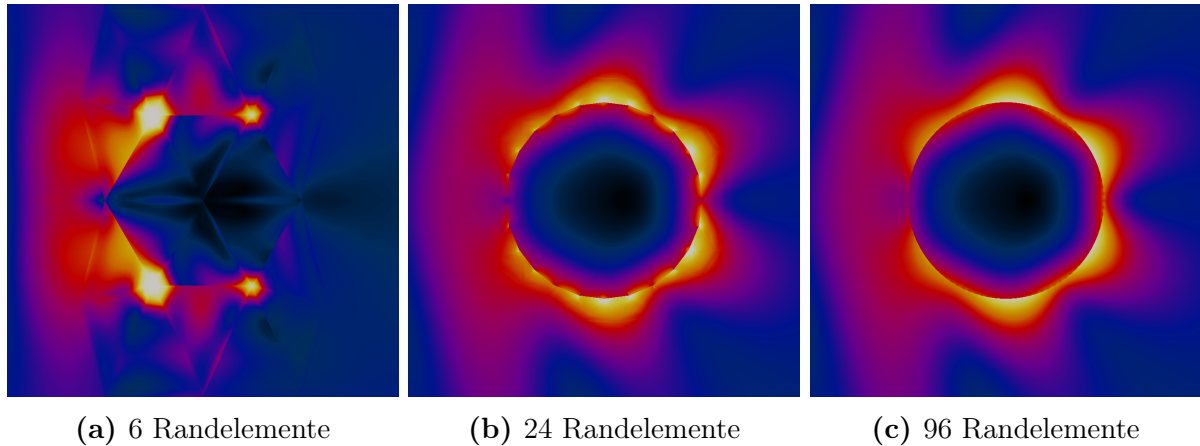
Vergleicht man nun runde mit linearen Elementen, fallen zwei Dinge sofort auf: erstens, sind bei gleicher Diskretisierung die Fehler der runden Elemente deutlich niedriger und zweitens ist das spektrale Konvergenzverhalten gleichmäßiger. Dabei ist analog zu den Betrachtungen im Abschnitt 4.2 das Netz mit sechs runden Oberflächenelementen etwa genauso gut wie jenes mit 24 linearen Oberflächenelementen, so dass sich mit einem deutlich gröberen Netz ähnlich gute Ergebnisse erzielen lassen, was wiederum Rechenressourcen schont.

Nach den vorgestellten integraldominierten Messgrößen lohnt sich, speziell bei der plasmonischen Nanooptik, noch ein Blick auf die Nahfeldüberhöhung und damit eine sehr sensitive punktuelle Messgröße. Die maximale Nahfeldüberhöhung ist für ausgewählte Rechnungen in Abbildung 5.1.6 dargestellt. Hier wird der Hauptunterschied zwischen runden Elementen und damit geometrisch genauerer Repräsentation des Systems und linearen Elementen sichtbar. Während alle Ergebnisse mit runden Elementen sehr schnell konvergieren und selbst für das gröbste Netz in der richtigen Größenordnung liegen, benötigen die linearen Elemente das feinste Netz um die richtige Dimension der Nahfeldüberhöhung vorherzusagen. Die Ursache liegt in der, durch die Diskretisierung hervorgerufenen, Approximation der Geometrie, die dazu führt, dass eine Kugel durch ein  $n$ -Eck repräsentiert wird. An den Ecken dieses der Simulation zugrunde liegenden  $n$ -Eckes treten nun Feldsingularitäten auf, welche als Eckplasmonen bezeichnet werden können. Besonders eindrucksvoll ist

der Unterschied zwischen den runden und den linearen Elementen beim Vergleich der Feldbilder (Abbildungen 5.1.7 und 5.1.8) bei einer Wellenlänge von 208 nm zu sehen. Bei den

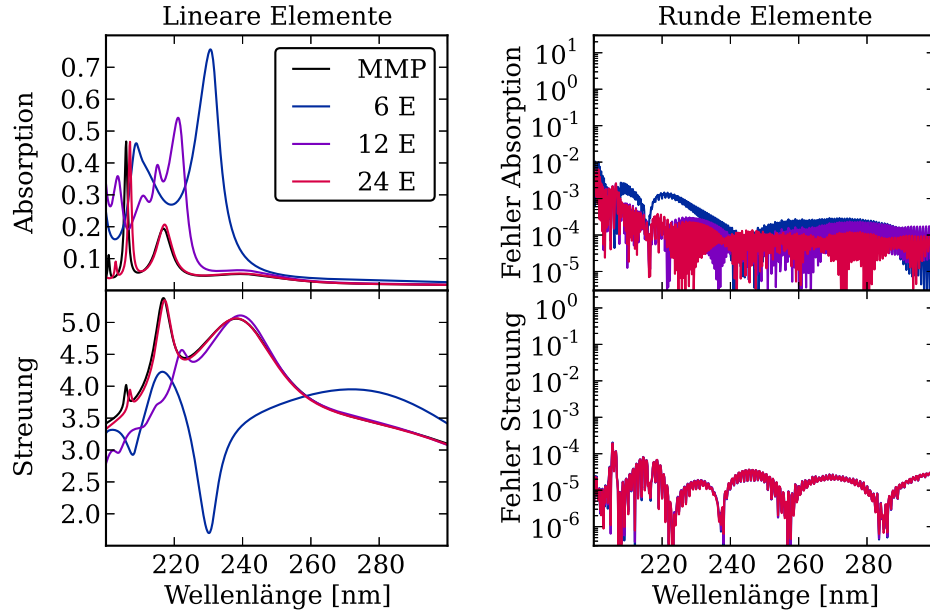


**Abbildung 5.1.7.:** Feldüberhöhung für runde Elemente mit uniformer Verfeinerung bei 208 nm Anregungswellenlänge, Bilder skaliert von 0 bis  $4.3E_0$ , Bildausschnitt  $300 \times 300$  nm.



**Abbildung 5.1.8.:** Feldüberhöhung für lineare Elemente selektive Verfeinerung bei 208 nm Anregungswellenlänge, Bilder skaliert von 0 bis  $4.3E_0$ , Bildausschnitt  $300 \times 300$  nm.

runden Elementen gibt es keine Diskretisierungsartefakte auf der Oberfläche des Partikels. Nur bei sechs Oberflächenelementen ist das Netz so grob, dass der exponentielle Abfall des Nahfeldes nicht gut genug modelliert werden kann. Bei den anderen Netzen mit runden Elementen ist zumindest optisch kein Unterschied mehr auszumachen. Das gegenteilige Bild bieten, wie erwartet, die Feldplots mit den linearen Elementen. Deutlich treten die, hier unerwünschten, Feldüberhöhungen als Folge der ungenügenden Oberflächendiskretisierung hervor. Allein für das feinste Netz sind diese „Hot-Spots“ so fein, dass sie zumindest

**Selektive Verfeinerung 8. Ordnung**

**Abbildung 5.1.9.:** Absorptions- und Streuspektren für lineare Elemente sowie die relativen Fehler für runde Elemente

optisch nicht mehr wahrgenommen werden können. Eine Folge dieser „Hot-Spots“ ist die ungenügende Feldapproximation, da an diesen Stellen das Feld der Theorie nach unendlich hoch sein sollte (siehe [1, 12, 41]), was mit der gewählten Basis nur schlecht gelingt. Eine weitere Folge dieser schlechten Approximation ist die erhöhte Absorption aufgrund der Dissipation im numerischen Fluss. Daher liegen die Streuspektren, insbesondere für das Netz mit 12-Randelementen, näher an der Referenzlösung als die Absorptionsspektren, die neben der fehlerhaften Geometrie auch noch die zusätzliche Dissipation des Systems beinhalten.

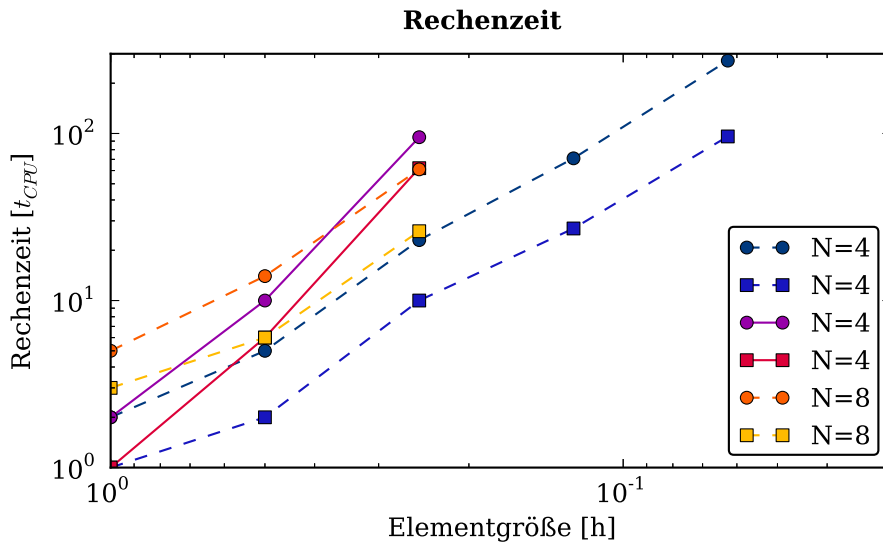
### Verhalten bei höherer Ordnung der Basisfunktionen

Zum Schluss dieses Abschnittes soll noch kurz der Verlauf bei weiterer Erhöhung der Approximationsordnung der Basispolynome dargelegt werden. Dazu sollen hier einerseits die Spektren für lineare Elemente und andererseits der spektrale Fehlerverlauf für die runden Elemente genügen. Beide sind in Abbildung 5.1.9 (vergleiche mit Abbildung 5.1.4 und 5.1.5) dargestellt. Betrachtet man die Ergebnisse für die linearen Elemente, so fällt auf, dass sich das Spektrum nicht in Richtung „2D-Kugel“ verschiebt, sondern der ursprüngliche Verlauf sich verfestigt, da nur das jeweilige  $n$ -Eck besser modelliert aber eben keine „Kugel“

beschrieben wird<sup>3</sup>. Dem gegenüber erreicht man mit runden Elementen sehr schnell die durch die anderen Systemkomponenten (PML, Konvergenz der Filter, etc.) vorgegebene Grenze, so dass weder eine weitere Erhöhung der Ordnung, noch eine feinere Diskretisierung zu einer deutlichen Verbesserung des Resultates führen. So ist in dem gewählten Beispiel der minimale Fehler bereits mit dem Netz, welches sechs Oberflächenelemente enthält, erreicht.

### Bemerkung zur Rechenzeit

Neben den qualitativen Verbesserungen, die sich mit den runden Elementen erzielen lassen, ist es für die Numerik auch entscheidend, ob sich damit die Rechenzeiten verkürzen lassen. Dazu wird wieder Gleichung 4.1.15 verwendet. Hier ist  $c_{Nt} \approx 1.4$  und nicht mehr Eins (vergleiche Kapitel 4.1 und 4.3), da nun neben den konvexen Verformungen auch konkave Elemente vorkommen, die eine Reduzierung des Zeitschrittes mit sich bringen. Des Weiteren lag der empirisch ermittelte Wert für die zusätzliche Rechenzeit je rundem Element bei  $c_{Nce} \approx 4.6$ , das heißt, dass ein rundes Element ca. 5.6 mal mehr Rechenzeit je Zeitschritt beansprucht als ein lineares Element. Die Rechenzeiten sind mit den so ermittelten Werten in Abbildung 5.1.10 zusammengefasst. Wie zu erwarten war, ist die Rechenzeit der



**Abbildung 5.1.10.:** Rechenzeiten für die 2D-Kugel. Kreise: runde Elemente, Vierecke: lineare Elemente, Linien: uniforme Verfeinerung, Striche: selektive Verfeinerung.

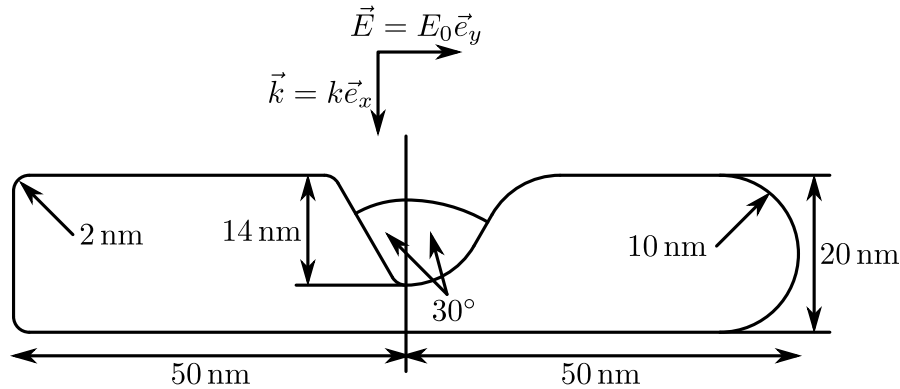
runden Elemente, bei gleichem Netz, größer als für lineare Elemente. Auch ist der Abstand

<sup>3</sup>Für  $\lim_{n \rightarrow \infty}$  kommt man der 2D-Kugel natürlich beliebig nahe. So zeigt in Abbildung 5.1.8 das 96-Eck bereits eine sehr gute Übereinstimmung mit dem Bild einer 2D-Kugel.

zwischen den Rechenzeiten für runde beziehungsweise lineare Elemente für ein selektiv verfeinertes Netz größer als für ein uniform verfeinertes Netz. Die letzte Beobachtung ist eine Folge des unterschiedlichen Anteils an runden zu linearen Elementen in den verschiedenen Netzen. So gibt es in einem selektiv verfeinerten Netz relativ gesehen mehr runde Elemente bezogen auf die Gesamtzahl an Elementen als in einem uniform verfeinerten Netz. Vergleicht man zuerst die Rechnung „achte-Ordnung-runde-Elemente-Sechseck-uniforme-Verfeinerung“ mit „vierte-Ordnung-runde-Elemente-12-Eck-uniforme-Verfeinerung“, da beide ungefähr den gleichen Fehler besitzen, so ist die höhere Ordnung ca. 2 mal schneller als die niedrigere, so dass es effektiver ist, mit höheren Ordnungen zu rechnen, wo dies sinnvoll ist. Um nun die Effizienz der runden mit derjenigen der linearen Elementen vergleichen zu können, wird hier die Rechnung „vierte-Ordnung-lineare-Elemente-48-Eck-selektive-Verfeinerung“ mit der Rechnung „vierte-Ordnung-runde-Elemente-12-Eck-selektive-Verfeinerung“ verglichen. Auch hier zeigen beide Rechnungen den ungefähr gleichen Fehler im Fernfeldverhalten (Absorption und Streuung). Dabei stellt sich heraus, dass die runden Elemente ca. 5.5 mal schneller sind als die linearen Elemente. Vergleicht man darüber hinaus auch das Nahfeld (gleiche Rechnung runde Elemente, mit dem 96-Eck-lineare-Elemente), so sind die runden Elemente sogar 19 mal schneller. Bei diesem Rechenzeitvergleich wurden die Frequenzfilter abgeschaltet, da sie in der hier verwendeten Implementation einen signifikanten Anteil an der Gesamtrechenzeit beanspruchen. Werden diese mit berücksichtigt ergibt sich ein weiterer Geschwindigkeitsvorteil zugunsten der runden Elemente, da einerseits weniger Elemente gefiltert werden müssen und andererseits der Anteil der runden Elemente an der Gesamtrechenzeit weiter sinkt.

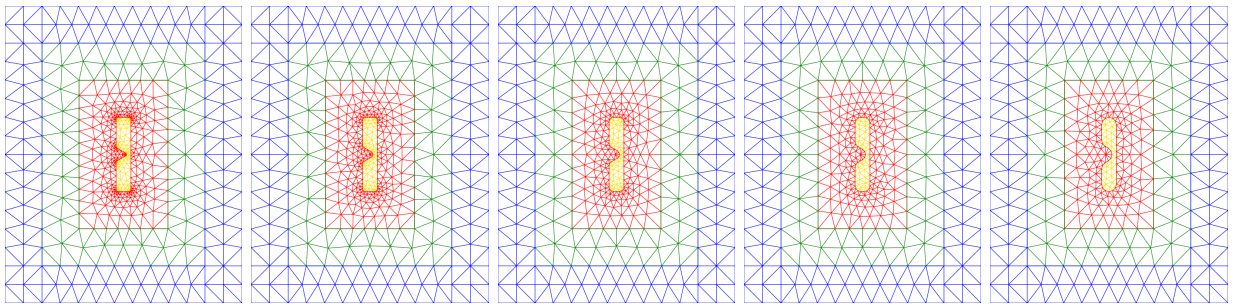
### 5.1.2. V-Graben

Nachdem im letzten Abschnitt ein eher akademisches Beispiel umfassend behandelt wurde, soll es im Folgenden um ein anwendungsorientiertes plasmonisches System, den V-Graben (engl.: V-Groove) gehen. Für diesen existiert im Allgemeinen keine geschlossene analytische Lösung mehr. Das Aussehen des hier simulierten Systems ist in Abbildung 5.1.11 skizziert. Das System besteht aus einem 100 nm breiten und 20 nm dicken Silberstreifen<sup>4</sup>, in dem



**Abbildung 5.1.11.:** Skizze eines V-Grabens mit variablem Rundungsradius. Links ist die Variante mit 2 nm und rechts die Variante mit 10 nm Rundungsradius skizziert.

sich ein V-förmiger Graben mit einer Tiefe von 10 nm einem halben Öffnungswinkel von 30° mit verschiedenen Ausrundungsradien von 2 nm bis 10 nm befindet. Die verwendeten Netze sind in Abbildung 5.1.12 dargestellt. Deutlich ist hier der große Dynamikumfang hin-

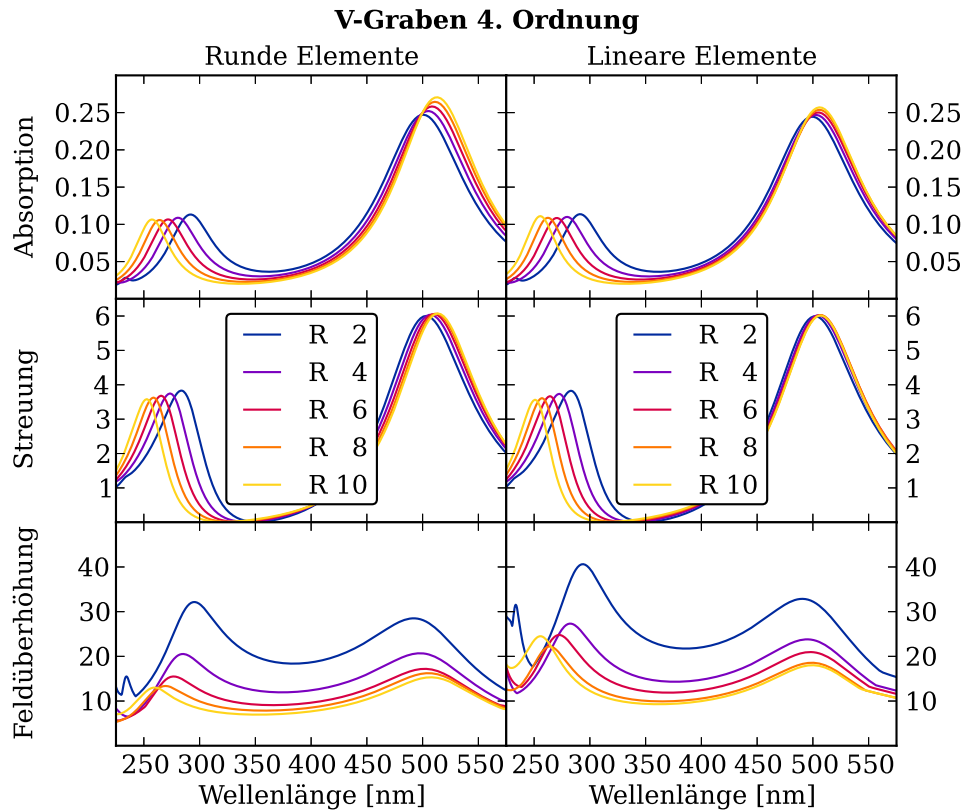


(a) 2 nm Radius, (b) 4 nm Radius, (c) 6 nm Radius, (d) 8 nm Radius, (e) 10 nm Radius, 1050 Elemente, davon 40 rund, 910 Elemente, davon 40 rund, 794 Elemente, davon 40 rund, 738 Elemente, davon 40 rund, 672 Elemente, davon 40 rund

**Abbildung 5.1.12.:** Serie der V-Graben-Netze mit variablen Rundungsradien. Die Farben grenzen die verschiedenen Simulationsgebiete (Gelb: Partikel, Rot: TF, Grün: SF und Blau: PML-Bereich) voneinander ab.

<sup>4</sup>Dabei wurde wieder das DRUDE-Silber aus Tabelle 2.3.1 verwendet.

sichtlich der Elementgröße zu sehen. So hat das kleinste Element eine Kantenlänge von ca.  $\approx 1$  nm, während die größten Elemente Kantenlängen von  $\approx 25$  nm aufweisen<sup>5</sup>. In jedem Netz gibt es 40 gekrümmte Elemente. Die Anregung erfolgte wieder mittels eines differenzierten GAUSSschen Pulses mit einer Polarisierung entlang der Hauptachse des V-Grabens, das heißt,  $\vec{E} = E_0 \vec{e}_y$ . Die erhaltenen Absorptions-, Streu- und Feldüberhöhungsspektren sind in Abbildung 5.1.13 jeweils für Rechnungen mit linearen und runden Elementen, beide mit einer polynomialen Ordnung  $N = 4$  ermittelt, gezeigt. Zuerst fallen die beiden



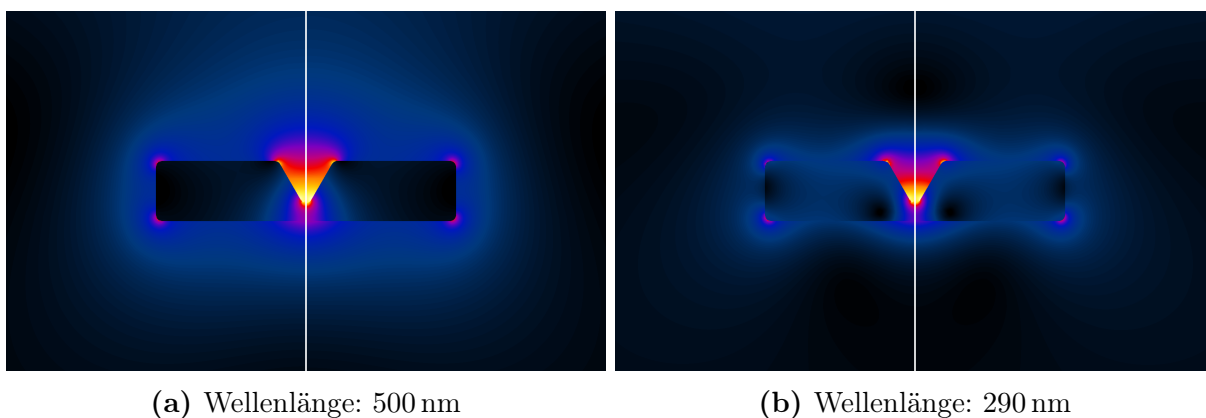
**Abbildung 5.1.13.:** Absorptions- und Streuspektren, sowie Feldüberhöhung für runde und lineare Elemente. Die Variable  $R$  gibt den Rundungsradius der Ecken des V-Grabens in nm an.

„Hauptresonanzen“ des Systems, bei ca. 275 nm und ca. 500 nm, auf. Diese beziehen sich auf die Dipolresonanz (ca. 500 nm) des 100 nm langen Streifens, sowie der gekoppelten Resonanz (ca. 275 nm) der beiden 50 nm kurzen, durch den Graben geteilten, Segmente. Des Weiteren stellt man fest, dass die Verschiebung der Resonanzen für die runden Elemente stärker ist als für die linearen Elemente.

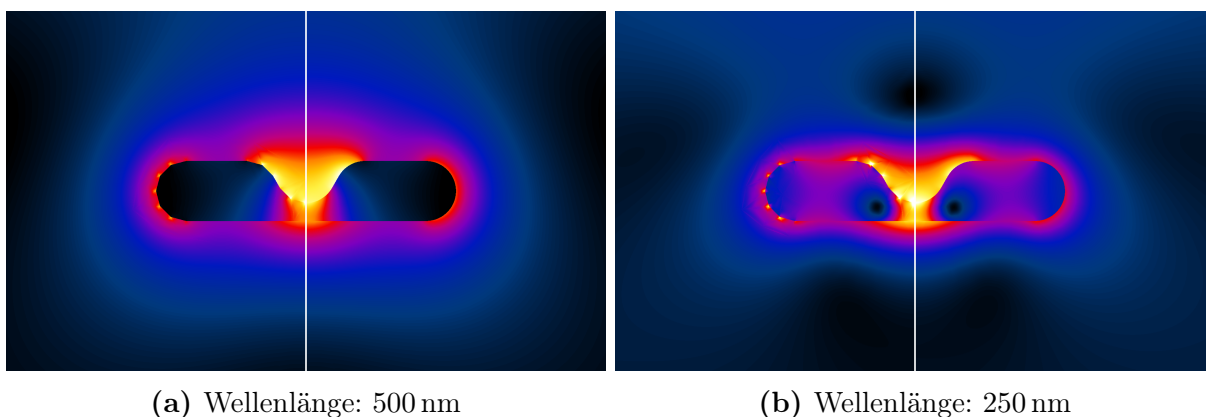
<sup>5</sup>Zum Vergleich: bei dem feinsten Netz aus Abschnitt 5.1.1, siehe Abbildung 5.1.3d, hatte das kleinste Element eine Kantenlänge von ca. 4.9 nm.



Darüber hinaus ist die Resonanz bei ca. 275 nm deutlich sensitiver auf den Ausrundungsradius, was sich anhand der stärker radiusabhängigen Maximaposition bemerkbar macht. Die Ursache liegt darin begründet, dass ein kleinerer Radius die kurzen Segmente erstens länger macht und zweitens näher zueinander rückt. Beide Effekte bewirken eine Rotverschiebung. Auch bei diesem System liegt die Vorhersage der Feldüberhöhung für die linearen Elemente etwas (ca. 1.2-1.8 mal) höher als für die runden Elemente. Dieser Effekt ist für kürzere Wellenlängen stärker ausgeprägt. Obwohl die Wellenlänge deutlich größer als der Rundungsradius ( $\lambda/r > 100$ ) ist, kann sie selbst bei dem kleinsten Rundungsradius noch festgestellt werden. Nichtsdestotrotz geben auch die linearen Elemente sowohl den Verlauf, als auch die Größenordnung des Feldes korrekt wieder. Um die Unterschiede zwi-

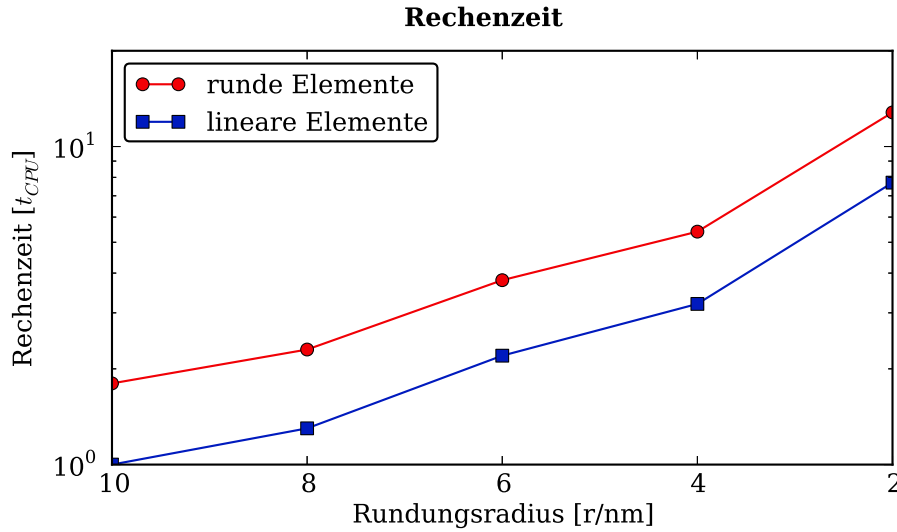


**Abbildung 5.1.14.:** Gegenüberstellung der Feldüberhöhung für runde und lineare Elemente bei einem Eckradius von 2 nm. Jeweils linke Bildhälfte: lineare Elemente, recht Bildhälfte: runde Elemente.



**Abbildung 5.1.15.:** Gegenüberstellung der Feldüberhöhung für runde und lineare Elemente bei einem Eckradius von 10 nm. Jeweils linke Bildhälfte: lineare Elemente, rechte Bildhälfte: runde Elemente.

schen den runden und linearen Elementen genauer beleuchten zu können, kann man sich



**Abbildung 5.1.16.:** Rechenzeiten für den 2D-V-Graben. Da hier der relative Anteil der runden Elemente am gesamten Netz gering ist, sinkt das Verhältnis der Rechenzeit zwischen runden und linearen Elementen auf  $c_{Nt} = 1.4$ , den Wert der, aufgrund der runden Elemente, erweiterten Zeitschrittanzahl.

analog zur 2D-Kugel (siehe Abschnitt 5.1.1) wieder die Feldplots an den Resonanzen ansehen. Diese sind als Gegenüberstellung von runden und linearen Elementen in Abbildung 5.1.14 für den kleinsten Radius von 2 nm und in Abbildung 5.1.15 für den größten Radius von 10 nm dargestellt. Dabei fällt im Vergleich der beiden Radien auf, dass beim kleinsten die Feldüberhöhung an den Ecken der linearen Elemente fast nicht mehr wahrnehmbar ist, während beim größten noch deutliche „Hot-Spots“ zu erkennen sind. Dies führt dazu, dass beim 10 nm Radius nicht nur in der näheren Umgebung der kritischen „Ecken“, sondern auch in entfernterem Bereich Unterschiede zwischen den linearen und den runden Elementen feststellbar sind.

Darüber hinaus ist das Feld beim 10 nm Radius eher um den ganzen V-Graben verteilt, während es sich beim 2 nm Radius auf den eigentlichen Graben und die Ecken des Systems beschränkt.

#### Anmerkung zur Rechenzeit

Analog zum vorangegangenen Abschnitt 5.1.1 sind auch hier zwei Anteile wesentlich, die die Rechenzeit der runden Elemente vergrößert: Einerseits der zusätzliche Aufwand, der für jedes runde Element notwendig ist ( $c_{Nce} \approx 5$ ), und andererseits die Verringerung des kritischen Zeitschrittes ( $c_{Nt} \approx 1.4$ ). Die normierten Rechenzeiten sind in Abhängigkeit des Rundungsradiuses in Abbildung 5.1.16 dargestellt. Da hier der Anteil der runden Elemente

am Gesamtnetz vernachlässigbar ist (er liegt zwischen 3.8 % und 5.9 %) ist der verringerte kritische Zeitschritt die Hauptursache für die längere Rechenzeit der runden Elemente. Dies ist deutlich am ungefähr gleichen Verhältnis der Rechenzeiten zwischen runden und linearen Elementen bei gleichem System zu sehen. Darüber hinaus wird gut der unterschiedliche Einfluss von Elementgröße und Elementanzahl verdeutlicht. So nimmt die Elementanzahl vom größten zum kleinsten Radius nur um den Faktor zwei zu, gleichzeitig nimmt die Größe des kleinsten Elementes um den Faktor fünf ab. Dies führt letztendlich zu einer verlängerten Rechenzeit um den Faktor<sup>6</sup>  $2 \cdot 5 = 10$ .

---

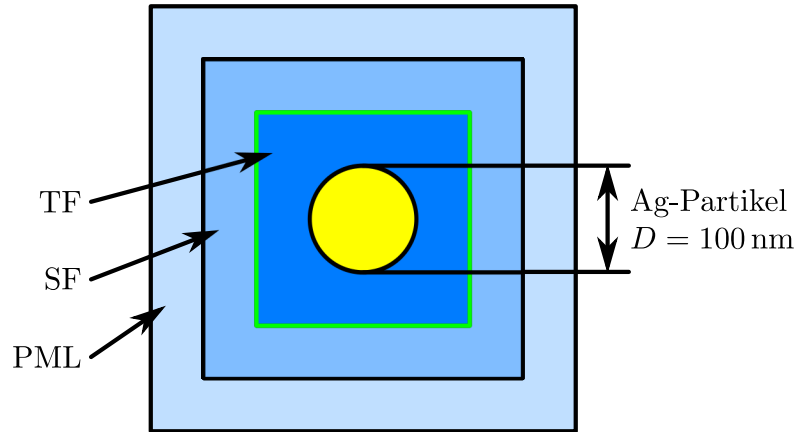
<sup>6</sup>Nach den Betrachtungen aus Abschnitt 3.1.3 geht die Größe des kleinsten Elementes invers in die Rechenzeit ein.

## 5.2. 3D-Plasmonische Nanostrukturen

Obgleich sich zweidimensionale Systemen deutlich einfacher berechnen und auch hinsichtlich der Netze und Felder besser darstellen lassen, sind die meisten praktischen Probleme dreidimensionaler Natur. Daher soll hier im Folgenden der dreidimensionale Fall behandelt werden. Begonnen wird wieder mit einem Referenzsystem, der 3D-Kugel, bevor zwei weitere lineare Systeme, Patchantenne und s-SNOM, behandelt werden.

### 5.2.1. Partikel

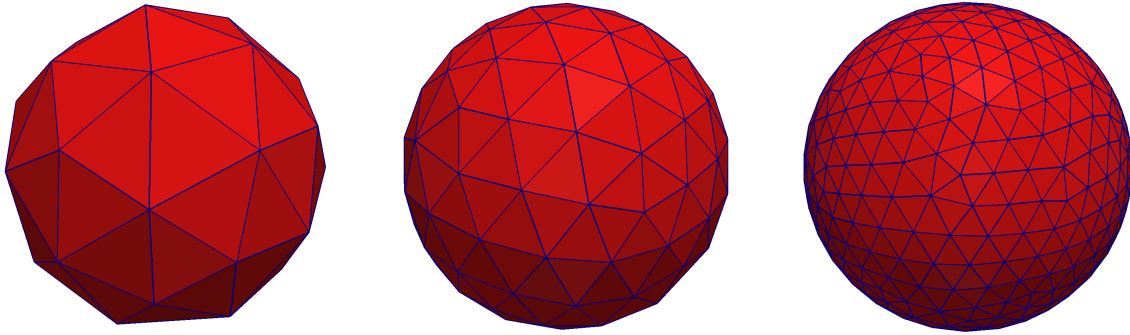
Analog zu Abschnitt 5.1.1 wird hier eine, nun dreidimensionale, Silber-Kugel mit einem Durchmesser von 100 nm behandelt. Eine Skizze des Systems findet sich in Abbildung 5.2.1. Das Umgebungsmedium besteht aus Luft ( $n = 1.0$ ) und für das Silber wird wieder das



**Abbildung 5.2.1.:** Skizze der 3D-Kugel. Das System besteht aus insgesamt vier Bereichen: der 3D-Kugel, dem TF-Bereich, dem SF-Bereich und dem abschließendem PML

DRUDE-Modell aus Tabelle 2.3.1 angesetzt. Die Anregung besteht aus einem differenzierten GAUSSschen Puls mit der Mittenfrequenz von 320 nm. Das Frequenzspektrum wird im Bereich von 200 nm bis 800 nm in 200 äquidistanten Frequenzpunkten mittels eines DFT-Filters (siehe 2.6.4) aufgezeichnet. Ähnlich wie in Abschnitt 4.3 soll auch hier wieder der Einfluss der Genauigkeit der Oberflächenrepräsentation in Abhängigkeit von der Oberflächendiskretisierung betrachtet werden. Die Motivation folgt aus dem Umstand, dass der hier verwendete Netzgenerator NETGEN (siehe [60]) nur die Oberflächenapproximation bis zur Ordnung  $B = 2$  beherrscht. Zu diesem Zweck wurden drei verschiedene Netze mit einer Oberflächendiskretisierung von  $r/h_{min} = R = 1.5, 3.0, 6.0$  erstellt. Hier ist  $r = 50 \text{ nm}$  der Radius der Kugel und  $R$  die Anzahl an Oberflächenelementen je Kugelradius, das heißt, ein Maß für die Diskretisierung der Oberfläche. Die Kugeloberfläche ist in Abbildung

5.2.2 dargestellt. Die Knotenpunkte auf der Oberfläche wurden, ausgehend von den vom



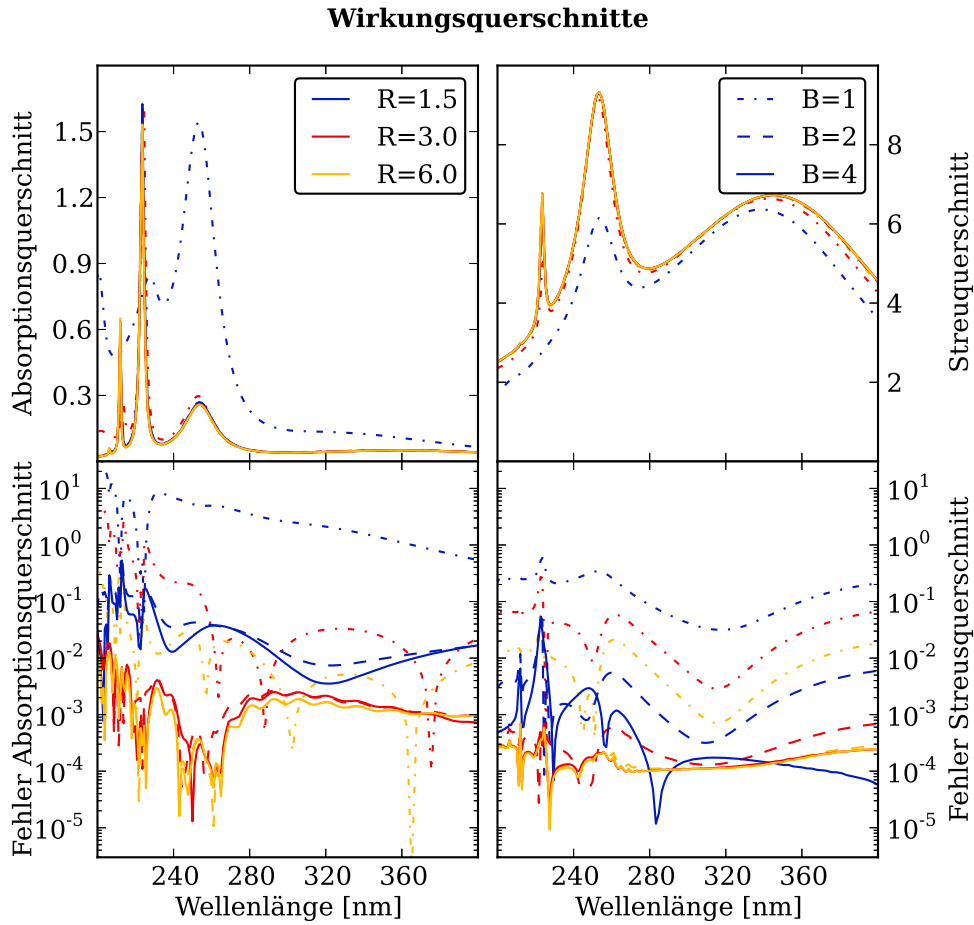
(a) Kugeloberfläche des Netzes mit  $h_{min} \approx r/1.5$  (insgesamt 3440 Elemente, 254 davon rund) (b) Kugeloberfläche des Netzes mit  $h_{min} \approx r/3.0$  (insgesamt 4617 Elemente, 765 davon rund) (c) Kugeloberfläche des Netzes mit  $h_{min} \approx r/6.0$  (insgesamt 11170 Elemente, 3008 davon rund)

**Abbildung 5.2.2.:** Kugeloberflächen aus einer Serie selektiv verfeinerter Netze. Bei der Zählung der runden Elemente wurden alle Elemente gezählt, welche gekrümmt werden mussten, das heißt, inklusive der Elemente, welche nur über eine Kante auf der Oberfläche verfügen und der Elemente, welche im TF-Gebiet liegen.

Netzgenerator gelieferten Netzen, mittels des in Abschnitt A.2 beschriebenen Algorithmus bestimmt. Damit sind auch Oberflächenapproximationsordnungen mit  $B > 2$  möglich. Die Approximationsordnung der Felder war  $N = 4$ . Als Referenzmethode zur Bestimmung der Fehler für die Absorptions- und der Streuquerschnitte dient hier wieder MMP<sup>7</sup>. Zunächst kann man auch hier wieder die integralen Werte, das heißt Streu- und Absorptionsspektrum, betrachten. Diese sind in Abbildung 5.2.3 dargestellt. Wie erwartet, sind die linearen Elemente ( $B = 1$ ) schlechter als die runden Elemente ( $B > 1$ ). Darüber hinaus fällt der im Vergleich zu den linearen Elementen deutlich kleinere Abstand zwischen quadratischer Oberflächenapproximation ( $B = 2$ ) und voller, das heißt  $N = B$ , Oberflächenapproximation<sup>8</sup> ( $B = 4$ ) auf. Besonders wird dies anhand des Absorptionsquerschnittes deutlich, wo beide ein ähnliches Fehlerverhalten zeigen. Beim Streuquerschnitt ist der Unterschied zwischen  $B = 2$  und  $B = 4$  im Bereich der niedrigen Resonanzen, der Dipolresonanz bei ca. 345 nm und der Quadrupolresonanz bei ca. 250 nm, zwar etwas größer, für praktische Anwendungen aber immer noch hinreichend klein. So lassen sich die Unterschiede nur im Fehlerbild ausmachen. Anhand der Wirkungsquerschnitte kann man bereits das kleinste

<sup>7</sup>Die Ergebnisse wurden zudem mit einem hausinternen MIE-Programm nach [4] verifiziert.

<sup>8</sup>Da hier die Ordnung der Approximationspolynome maximal ausgenutzt wird. Eine Oberflächenrepräsentation  $B > N$  führt zu keiner weiteren Verbesserung, da die Approximationspolynome diese hohe Auflösung nicht abbilden können.

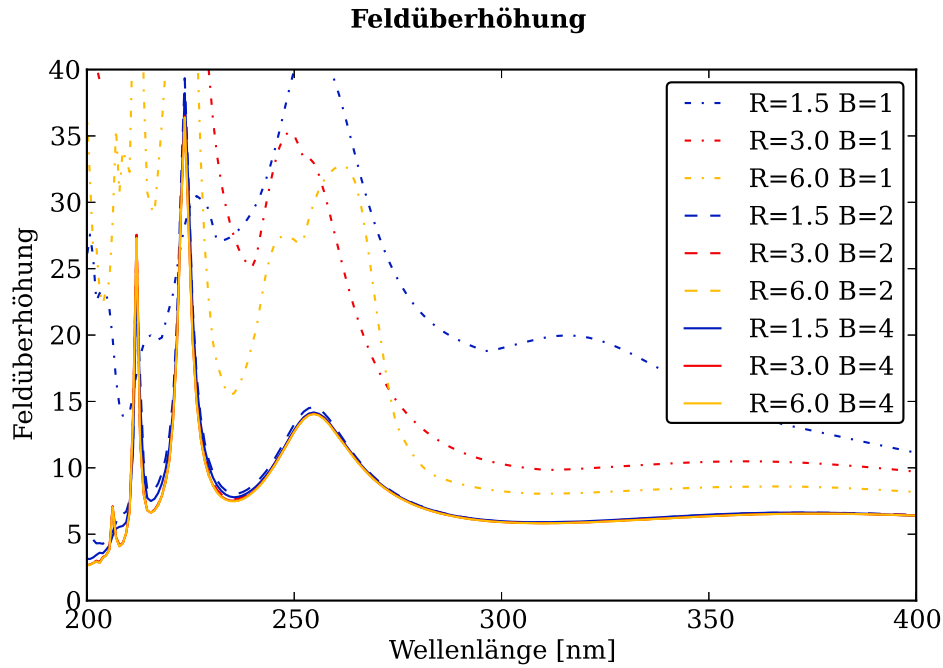


**Abbildung 5.2.3.:** Streu- und Absorptionsquerschnitte der 3D-Kugel sowie deren relative Fehler in Abhängigkeit der Oberflächenapproximation und -diskretisierung. Die Variable  $R = r/h_{min}$  ist (mit  $r$  der Radius der Kugel) ein Maß für die Oberflächendiskretisierung. Die Variable  $B$  beschreibt die Ordnung der Oberflächenrepräsentation ( $B = 1$ : lineare Elemente,  $B > 1$ : runde Elemente). Diese ist nicht zu verwechseln mit der Ordnung  $N$  der Feldapproximation.

Netz (5.2.2a) sowohl mit  $B = 2$  und  $B = 4$  als hinreichend nahe an der exakten Lösung ansehen. Beide Oberflächenapproximationsordnungen erreichen mit dem nächsten Netz (5.2.2b) den systembedingt kleinsten Fehler. Die Ursache liegt in der hier zur Anwendung gelangten selektiven Verfeinerung der Netze, die nur die Oberfläche betrifft, so dass mit fortschreitender Verfeinerung die Volumenfehler dominant werden. Dem gegenüber lässt sich bei den linearen Elementen ( $B = 1$ ) mit jeder Verfeinerung auch der Fehler weiter reduzieren. Nichtsdestotrotz kann auch hier mit dem mittleren Netz (5.2.2b) eine für die Praxis hinreichende Genauigkeit erzielt werden. Eine zu den runden Elementen äquivalente Konvergenz ist aber erst mit der feinsten Oberflächendiskretisierung (5.2.2c) zu erreichen.

Damit entspricht die Leistungsfähigkeit der runden und linearen Elementen im 3D ungefähr der im 2D. So entsprechen die drei Netze im Zweidimensionalen, mit der Nomenklatur von Abschnitt 5.1.1, vom größten zum feinsten Netz 9, 19 und 38 Randelementen.

Neben den Wirkungsquerschnitten kann man sich nun noch die maximale Feldüberhöhung anschauen. Diese ist in Abbildung 5.2.4 dargestellt. Auch hier zeigen die runden Ele-

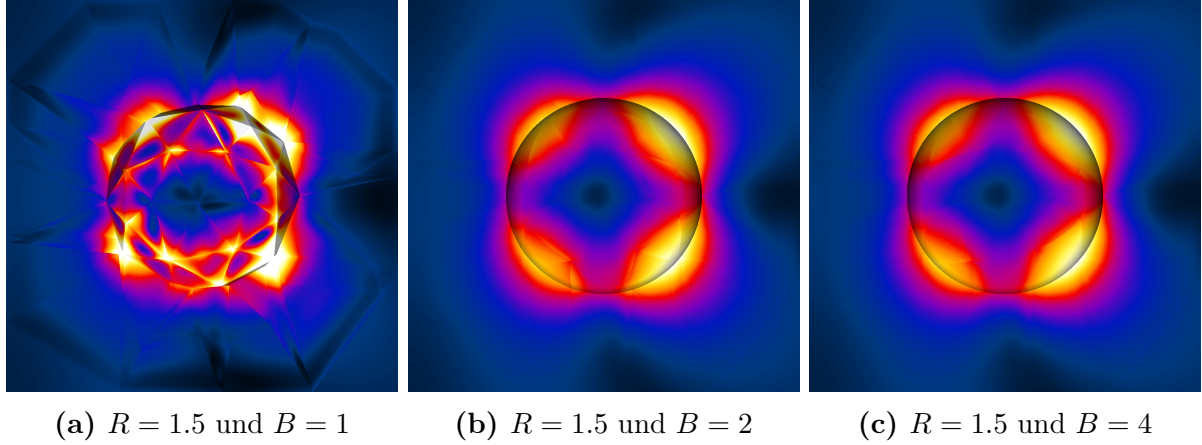


**Abbildung 5.2.4.:** Maximale Feldüberhöhung der 3D-Silber-Kugel. Die Variable  $R = r/h_{min}$  ist, mit  $r$  dem Radius der Kugel, ein Maß für die Oberflächendiskretisierung. Die Variable  $B$  beschreibt die Ordnung der Oberflächenrepräsentation ( $B = 1$ : lineare Elemente,  $B > 1$ : runde Elemente). Diese ist nicht zu verwechseln mit der Ordnung  $N$  der Feldapproximation!

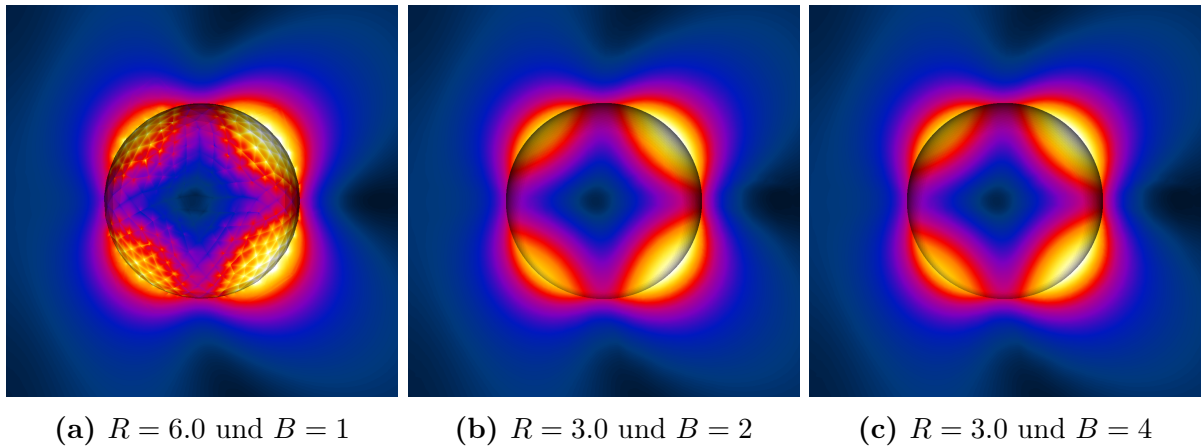
mente, sowohl in quadratischer wie auch voller Oberflächenapproximation, ein hervorragendes Verhalten. So gibt bereits das größte Netz (Abbildung 5.2.2a), mit runden Elementen gerechnet, alle spektralen Resonanzen sowohl in der Position wie auch der Größenordnung gut wieder. Zwischen  $B = 2$  und  $B = 4$  sind nur kleinere Abweichungen auszumachen, so liegt die mit „voller“ Oberflächenapproximation ermittelte maximale Feldüberhöhung näher an der konvergierten Lösung als für die quadratische Oberflächenapproximation. Für die feineren Netze (Abbildung 5.2.2b und Abbildung 5.2.2c) lässt sich, für  $B = 2$  und  $B = 4$ , im Spektrum kein Unterschied mehr ausmachen, so dass man das Ergebnis ab  $R = 3.0$  als konvergiert ansehen kann. Im Gegensatz dazu (und auch zum Zweidimensionalen) ließ sich hier mit den verwendeten Netzen und linearen Elementen keine auch nur annähernde



Konvergenz erreichen. So weichen alle Rechnungen sowohl in der Amplitude wie auch dem spektralen Verlauf erheblich von der mit den runden Elementen ermittelten Lösung ab. Abschließend seien noch die Feldbilder (gemitteltes  $E$ -Feld) für die Quadrupolresonanz bei ca. 224 nm gezeigt. Diese sind für alle Oberflächenapproximationsordnungen und das kleinste Netz in Abbildung 5.2.5 und für ausgewählte höhere Netze in Abbildung 5.2.6 dargestellt. Auch anhand dieser Feldbilder lässt sich, speziell für das größte Netz, die



**Abbildung 5.2.5.:** Nahfeld der 3D-Kugel bei ca. 224 nm für das größte Netz.



**Abbildung 5.2.6.:** Nahfeld der 3D-Kugel bei ca. 224 nm.

qualitative Verbesserung durch die runden Elemente visualisieren. So zeigen die linearen Elemente (Abbildung 5.2.5a) wieder ausgeprägte „Hot-Spots“, welche einerseits zu einem stark überschätzten Nahfeld führen und, aufgrund ihrer schlechten Approximation durch DG-Elemente, die Absorption, unter anderem durch Dissipation im numerischen Fluss, erhöhen. Dies erklärt auch, dass, im Vergleich zum Streuspektrum, ungenauer berechnete Absorptionsspektrum, welches neben den Geometriefehlern auch noch Dissipation des



DG-Systems beinhaltet. Im Vergleich der beiden runden-Elemente-Rechnungen ( $B = 2$  und  $B = 4$ ) beim niedrigsten Netz (Abbildung 5.2.5b und 5.2.5c) sind die Unstimmigkeiten schwerer auszumachen. So zeigt das Netz mit quadratischer Oberflächenapproximation zwar etwas mehr Unstetigkeiten als das Netz mit  $B = 4$ , dennoch erscheint es qualitativ deutlich näher an der exakten Lösung als die lineare Variante und zeigt praktisch keine „Hot-Spots“. Geht man nun noch ein Netz weiter, das heißt, runde Elemente und  $R = 3.0$  (Abbildung 5.2.2b), so sieht man ein störungsfreies Nahfeld ohne unphysikalische Artefakte. Dem gegenüber zeigt selbst das feinste Netz (Abbildung 5.2.2c), gerechnet mit linearen Elementen, obwohl es auch im Nahbereich (nächste Elementreihe nach den Oberflächenelementen) keine Artefakte mehr zeigt, immer noch „Hot-Spots“ im Nächstbereich (Gebiet um die Oberflächenelemente). Das heißt, mit feiner werdender Oberflächendiskretisierung lässt sich mit linearen Elementen der Unsicherheitsbereich um die Kugel zwar verringern, aber nicht gänzlich ausschließen.

### Bemerkung zur Rechenzeit

Analog zum 2D, speziell Abschnitt 5.1, ist auch hier wieder eine Zeitschrittanpassung aufgrund der konkaven Verformung der runden Elemente im TF-Bereich notwendig. Dieser Faktor liegt bei  $c_{Nt} = 1.2$ . Darüber hinaus wurde der zusätzliche Aufwand je rundes Element mit  $c_{Nce} = 4$  bestimmt. Aus diesen Werten und mittels Formel 4.1.15 lassen sich die normierten Rechenzeiten  $t_{CPU}$  ermitteln. Sie sind in Tabelle 5.2.1 dargestellt. Dabei gilt es zu bedenken, dass die Rechenzeit für ein rundes Element unabhängig von der Genauigkeit der Oberflächenrepräsentation ist, das heißt, dass die Rechenzeiten für  $B = 4$  und  $B = 2$  identisch sind. Neben der trivialen Erkenntnis, dass die Rechenzeiten für runde Elemente

**Tabelle 5.2.1.:** Normierte Rechenzeiten  $t_{CPU}$  für die 3D-Kugel

$R$	$B = 1$	$B > 1$
1.5	1.0	1.6
3.0	2.0	4.0
6.0	9.7	24.3

wieder größer sind als für lineare Elemente, fällt auf, dass der Abstand der Rechenzeiten mit feiner werdendem Netz deutlich steigt. So sind beim größten Netz die runden Elemente nur 1.6 mal langsamer als die linearen Elemente, während sie beim feinsten Netz schon 2.5 mal langsamer sind. Hier ist wieder die, relativ zu den linearen Elementen, wachsende Anzahl runder Elemente bei oberflächenselektiver Verfeinerung die Ursache. So sind

beim gröbsten Netz nur 7 % aller Elemente rund, während es beim feinsten Netz 27 % sind (vergleiche dazu die Angaben aus Abbildung 5.2.2). Vergleicht man nun die Rechenzeiten zwischen runden und linearen Elementen bei ungefähr gleichem Fehler, so sind die runden Elemente im Fernfeld ca. 6 mal schneller als lineare Elemente (feinstes Netz lineare Elemente versus gröbstes Netz runde Elemente). Kommt es nur auf ein grobes Spektrum an, so sind die runden Elemente bei ca. einer Größenordnung genauerem Ergebnis immer noch ca. 25 % schneller als die linearen Elemente (mittleres Netz lineare Elemente versus gröbstes Netz runde Elemente). Für das Nahfeld ist ein Vergleich der Rechenzeiten hier nicht möglich, da keine hinreichend genaue Lösung für die linearen Elemente existiert. Daher sind die runden Elemente, abhängig vom genauen Nahfeldbereich, ca.  $\geq 6$  mal schneller als die linearen Elemente. Damit ist für jeden Anwendungsfall die Benutzung der runden Elemente vorteilhaft.

### 5.2.2. Flache IR-Antennen

Ein Anwendungsgebiet der Plasmonik ist die oberflächenverstärkte Infrarotabsorptionsspektroskopie<sup>9</sup>. Ziel ist dabei, die Nahfeldverstärkung metallischer Nanostrukturen auf einem Substrat zu nutzen, um schwach absorbierende Moleküle zu detektieren [51]. Neben rauen Metallfilmen [48] und resonanten Gitterstrukturen [70] zeigen auch einfache Nanoantennen gut einstellbare Resonanzen [44]. Obgleich sich eine einzelne Antenne mittels elektrochemischer Synthese [25] herstellen, beziehungsweise Ensembles von diesen Antennen mittels fokussierter Ionenstrahlen<sup>10</sup> aus einer Metalloberfläche herausarbeiten lassen, sind beide Prozesse für die Herstellung großer Probenmengen ungeeignet, da bei den, häufig in Lösung hergestellten, Antennen das kontrollierte Aufbringen dieser auf ein Substrat schwierig und die FIB für einen großtechnischen Einsatz zu aufwendig ist. Einen Ausweg bilden Techniken, bei denen eine Schattenmaske, zum Beispiel mittels Elektronenstrahlolithographie, hergestellt wird, um dann die Metallstrukturen mit Hilfe dieser auf das Substrat aufzudampfen. Der Nachteil der Technik besteht darin, dass sich mit diesen Masken nicht beliebig kleine Strukturen herstellen lassen. Deshalb wird im Folgenden untersucht, wie sich die mittlere Intensitätsverstärkung  $\langle I \rangle$  einer Antenne auf einem Substrat als Funktion der Antennenbreite verhält. Dabei wird angenommen, dass sich um die Antenne nur eine Monolage von Molekülen befindet, die zudem nur schwach absorbiert, und damit die physikalischen Eigenschaften der Antenne nicht beeinflusst. Für die Simulation hat das den Vorteil, dass nur die Antenne, nicht jedoch die Molekülschicht, simuliert werden muss. Darüber hinaus lässt sich die Absorption nun vereinfacht mit Hilfe des Absorptionskoeffizienten  $\alpha$  und der Molekülschichtdicke  $\Delta z$  als Oberflächenintegral  $\int_S$  der Intensität  $I$  über die von Molekülen benetzte Antennenoberfläche  $S$  aufschreiben:

$$A \approx \alpha \Delta z \int_S I \quad (5.2.1)$$

Die Absorptionsverstärkung ist dann gegeben durch:

$$\frac{A}{A_0} \approx \frac{\int_S I}{\int_S I_0} \quad (5.2.2)$$

Diese Absorptionsverstärkung kann nun, da die Feldstärke  $E_0$  der Anregung und damit auch  $I_0 = |E_0|^2$  konstant sind, durch die gemittelte normierte Intensitätsverstärkung aus-

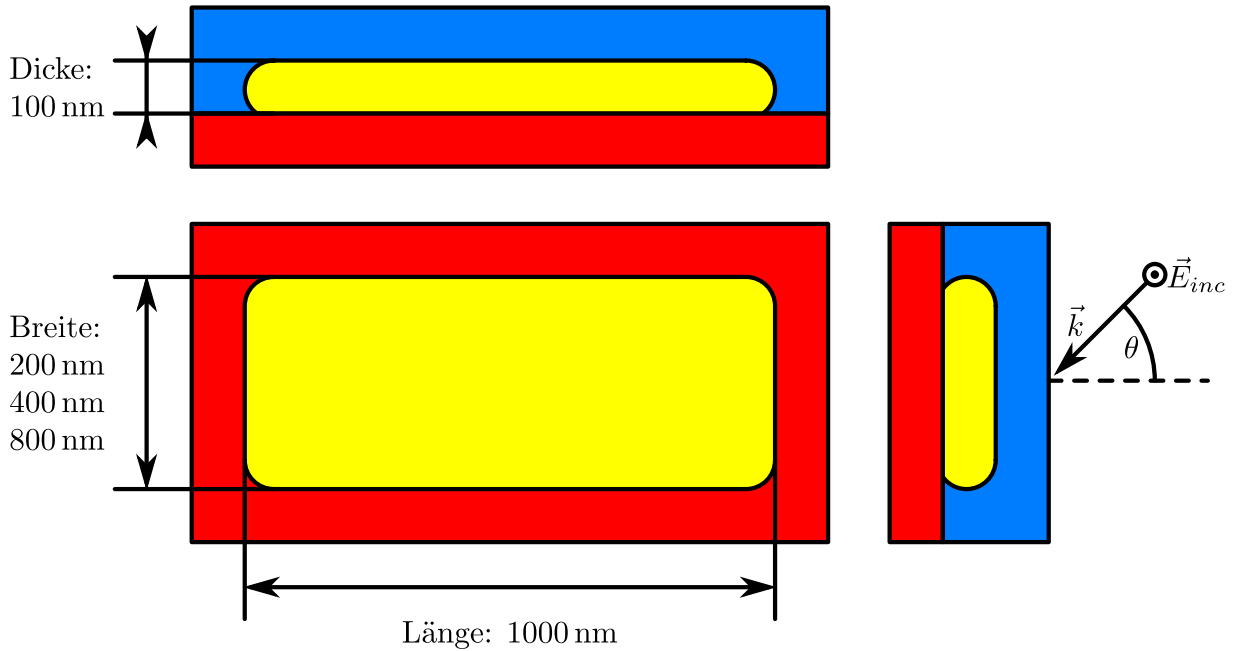
<sup>9</sup>engl: „Surface Enhanced InfraRed Absorption“ (SEIRA)

<sup>10</sup>engl: „Focused Ion Beam“ (FIB)

gedrückt werden:

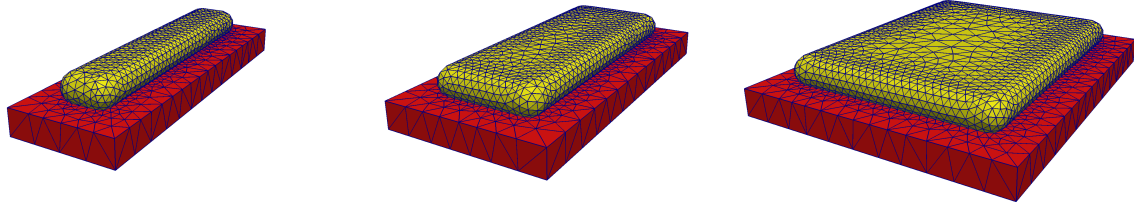
$$\frac{A}{A_0} = \frac{\langle I \rangle}{S} = \frac{\int_S \left| \frac{E}{E_0} \right|^2}{\int_S} \quad (5.2.3)$$

Dabei beschreibt  $S$  den Wert der Oberfläche der Antenne, die mit Molekülen benetzt ist. Die hier gewählten Abmessungen der Antenne an sich hat eine Länge von 1000 nm, eine Dicke von 100 nm und Breiten von 200, 400 und 800 nm. Als Materialmodell wurde DRUDE-LORENTZ-Gold (siehe Tabelle 2.3.1) verwendet. Die Antenne liegt auf einem Substrat mit einem Brechungsindex von  $n = 1.5$ . Neben der Breite der Antenne wurde auch der Einfallswinkel der Anregung, ein differenzierter GAUSSscher Puls mit der Mittenwellenlänge von  $2 \mu\text{m}$ , variiert. Eine Skizze des TF-Bereichs des Systems befindet sich in Abbildung 5.2.7. Die von Molekülen benetzte Oberfläche  $S$  ist durch die Gold-Luft-Grenze gegeben.



**Abbildung 5.2.7.:** Skizze der flachen Antenne. Gezeigt ist die Antenne (gelb), der TF-Bereich des Substrates (rot) und der TF-Bereich der Luft (blau). An diesen inneren Bereich schließt sich noch der 100 nm breite SF-Bereich und der 200 nm breite PML an (beide hier übersichtlichkeithalber weggelassen). Die Grenze zwischen Substrat und Luft wird dabei in den SF- und PML-Bereich fortgesetzt. Der Rundungsradius aller Antennenkanten ist 55 nm, das heißt, dass die Antennenunterkante mit dem Substrat einen Winkel von  $35.1^\circ$  einschließt.

Eine Darstellung der Netze der Antenne und des TF-Bereiches des Substrates befindet sich in Abbildung 5.2.8. Zum Vergleich wurden die gleichen Antennen jedoch ohne Substrat bei  $90^\circ$  Einfallswinkel berechnet. Bei der Ermittlung der Intensitätsverstärkung wurde dabei



(a) 200 nm breite Antenne (insgesamt 20990 Elemente, 3027 davon rund) (b) 400 nm breite Antenne (insgesamt 27589 Elemente, 3700 davon rund) (c) 800 nm breite Antenne (insgesamt 37336 Elemente, 4714 davon rund)

**Abbildung 5.2.8.:** Ausschnitt (Antenne und TF-Substrat) des Netzes für die Berechnung der flachen Antennen.

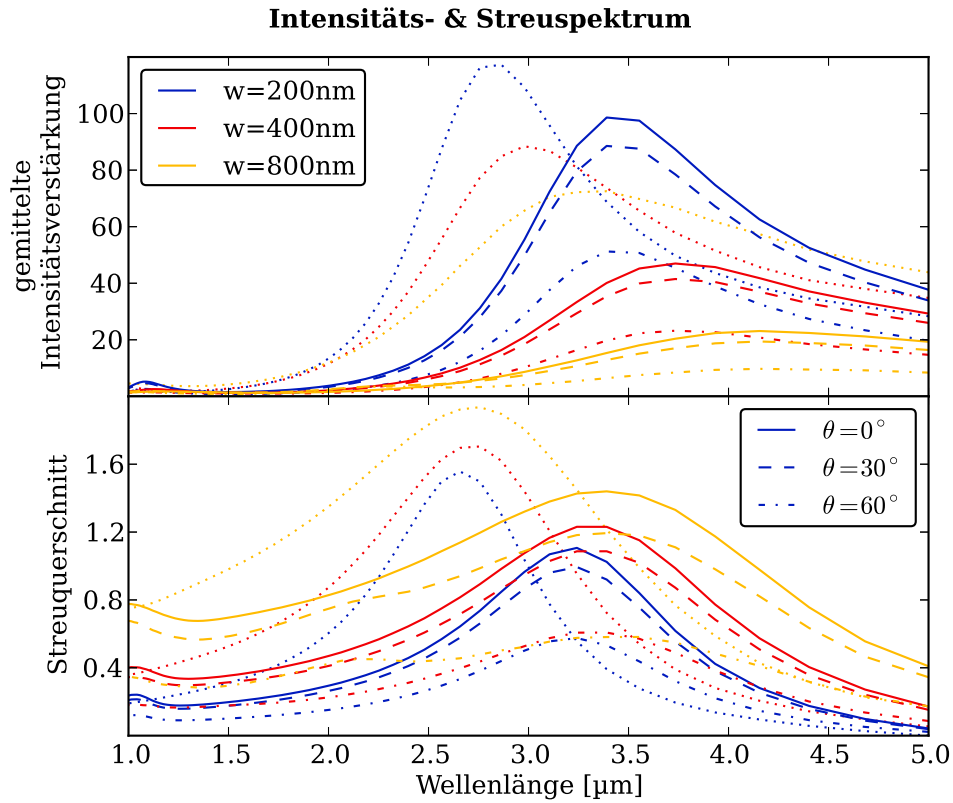
die gleiche Antennenoberfläche berücksichtigt, wie bei den Rechnungen mit Substrat. Die spektralen Eigenschaften wurden im Bereich von  $1\text{ }\mu\text{m}$  bis  $5\text{ }\mu\text{m}$  mit einem DFT-Filter an 80 Punkten bestimmt. Die mit der Formel 5.2.3 bestimmte Absorptionsverstärkung ist zusammen mit dem Streuquerschnitt<sup>11</sup> in Abbildung 5.2.9 dargestellt. Ganz allgemein stellt man fest, dass sich die Maxima für die Streuung und die gemittelte Intensitätsverstärkung mit breiter werdender Antenne zu größeren Wellenlängen hin verschieben. Dabei werden die Maxima einerseits breiter und andererseits für die Streuung höher und die gemittelte Intensitätsverteilung flacher. Für die Verschiebung sind hier zwei Ursachen verantwortlich. Die erste ist das Substrat, welches die Antenne in eine teilweise optisch dichtere Umgebung einbettet. Dieser Effekt liefert im Vergleich zu der 200 nm breiten Antenne in Luft, deren Maxima bei  $2.6\text{ }\mu\text{m}$  für die Streuung und bei  $2.8\text{ }\mu\text{m}$  für die Intensitätsverstärkung liegen, eine Rotverschiebung um  $0.8\text{ }\mu\text{m}$ . Die zweite Ursache ist das „offene“ Antennenende, welches für den Unterschied zwischen einer idealen Antenne, bei der die erste Resonanz bei dem Doppelten der Antennenlänge liegt, in diesem Fall also bei  $2L = 2\text{ }\mu\text{m}$ , und einer realen Nanoantenne verantwortlich ist [50, 6]. Der Einfallswinkel hat keinen Einfluss auf die Position, aber einen großen Einfluss auf die Höhe der Maxima, welche mit zunehmenden Einfallswinkel abnehmen. Die Ursache ist das mit dem Einfallswinkel abnehmende elektromagnetische Feld an der Luft-Substrat-Grenzfläche (mit  $r$  als Reflexionskoeffizient aus den FRESNELSchen Formeln 2.1.28):

$$E_{\text{Substrat}} = E_{\text{inc}} + E_{\text{ref}} \quad (5.2.4a)$$

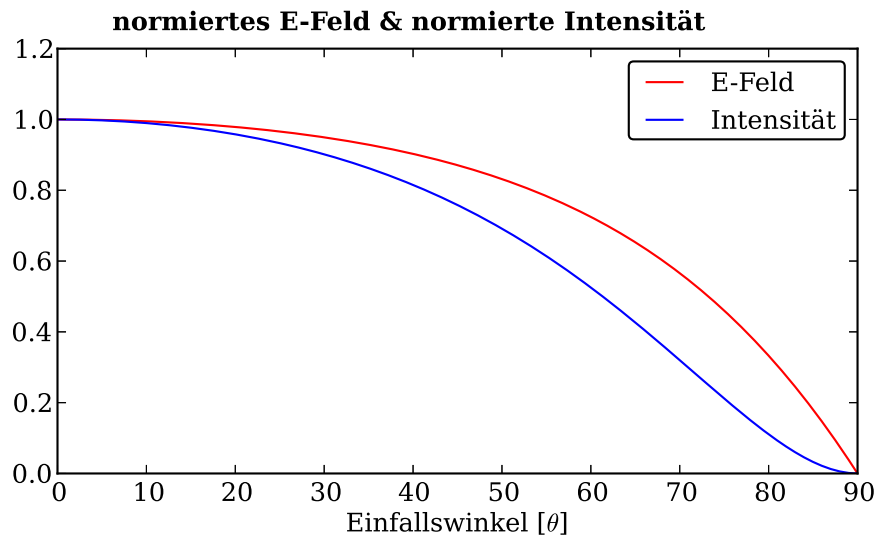
$$= 1 + r \quad (5.2.4b)$$

Der Verlauf ist in Abbildung 5.2.10 dargestellt.

<sup>11</sup>Der Streuquerschnitt wurde dabei als  $4\pi$ -Integral über alle Raumwinkel berechnet.



**Abbildung 5.2.9.:** Gemittelte Intensitäts- und Streuspektren der flachen Nanoantenne als Funktion der Antennenbreite und des Einfallswinkels. Die gepunkteten Kurven wurden zum Vergleich ohne Substrat bei  $90^\circ$  Einfallswinkel gerechnet.

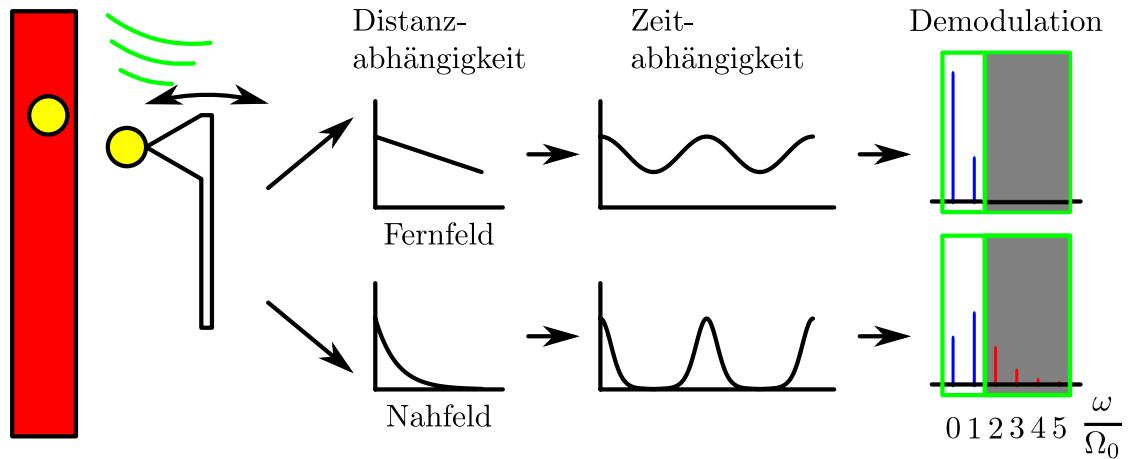


**Abbildung 5.2.10.:** Normierte Feldintensität und Feldstärke an der Luft-Glas-Grenzfläche in Abhängigkeit vom Einfallswinkel

Abschließend kann gesagt werden, dass eine Verbreiterung der Antenne für die Anwendung als absorptionsverstärkende Struktur kritisch zu sehen ist, da die Effizienz gegenüber einer schmalen Antenne um den Faktor Fünf schlechter ist. Zusätzlich ist fraglich, wie dicht diese Antennen auf einem Substrat aufgebracht werden können, ohne sich gegenseitig zu beeinflussen. Das heißt, für eine großflächig und günstig herzustellende absorptionsverstärkende Struktur sind breite Nanoantennen eher zweite Wahl.

### 5.2.3. s-SNOM

Bei der streuenden Nahfeldmikroskopie<sup>12</sup> wird mit einer beleuchteten Spitze eines Rasterkraftmikroskopes eine Probe abgetastet. Eine gute Übersicht zu diesem Thema findet sich in [23]. Durch die Beleuchtung der Spitze wird von dieser einerseits Licht gestreut und andererseits ein Nahfeld induziert. Nun kann sowohl das Streulicht, wie auch das Nahfeld mit der Probe wechselwirken, bevor es detektiert wird. Um nun beide Anteile zu trennen, wird die Spitze mit der Frequenz  $\Omega_0$  in Schwingung versetzt. Dadurch entsteht ein vom Proben-Spitzen-Abstand abhängiges Signal. Dabei ist das Streufeld linear und das Nahfeld nichtlinear vom Abstand abhängig. Zerlegt man nun das Messsignal, mit der Grundfrequenz  $\Omega_0$ , in seine FOURIERkomponenten, so hat das Streufeld nur Anteile bis zur Ordnung  $1\Omega_0$  und das Nahfeld Anteile in allen Ordnungen und speziell  $> 1\Omega_0$ . Ein Schema der Vorgehensweise findet sich in Abbildung 5.2.11. Mit Hilfe dieser Zerlegung können nun Materialwechsel-



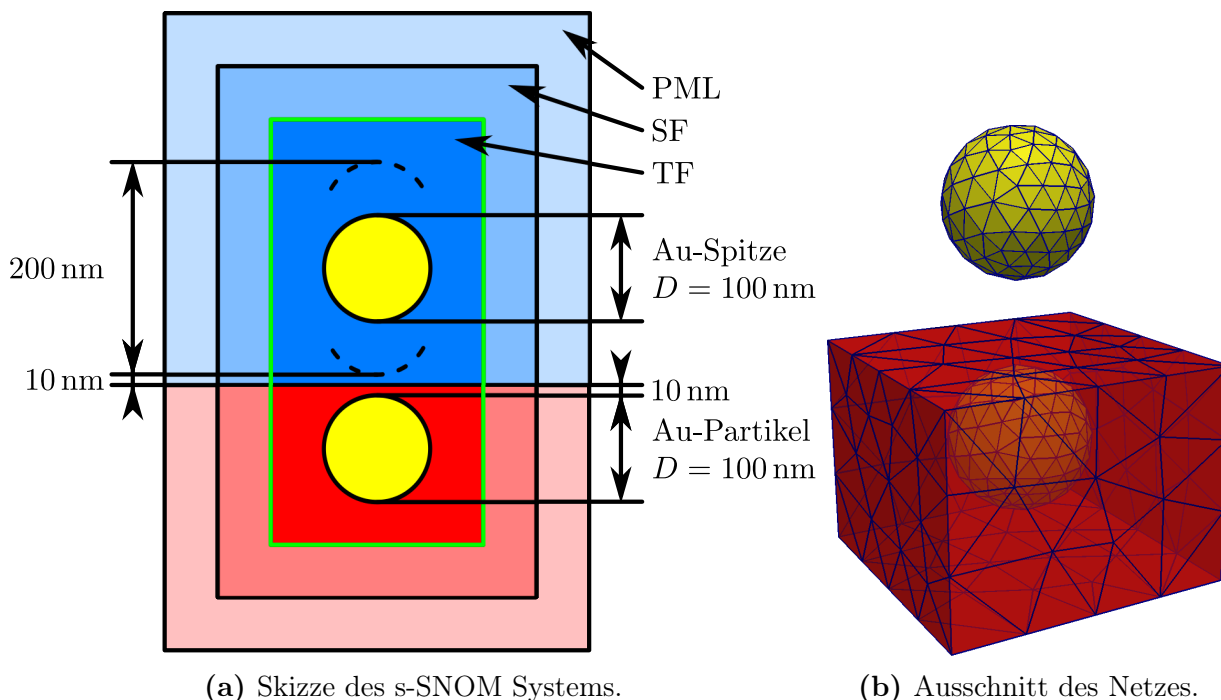
**Abbildung 5.2.11.:** Fern- und Nahfeld Demodulation beim s-SNOM nach [23]. Die unterschiedliche Abstandscharakteristik des Fern- und Nahfeldes, führen bei einer schwingenden Spitze zu verschiedenen spektralen Anteilen im zeitabhängigen Streusignal.

wirkungen getrennt von der Topographie untersucht werden. Da das Nahfeld eine deutlich kleinere Ausdehnung besitzt, typischerweise wenige 10 nm, als das Fernfeld, typischerweise  $> 0.61\lambda$  (mit  $\lambda$  der Wellenlänge des Lichtes), lassen sich mit dieser Technik Materialkontraste auf der sub- $\mu\text{m}$ -Skala sichtbar machen. Da für die Nahfeldwechselwirkung mit der Probe die Spitze eine entscheidende Rolle spielt, ist man dazu übergegangen, diese gezielt mit einem Metallnanopartikel zu funktionalisieren [69]. Für die weitere Behandlung wird daher die Spitze als „frei-schwebendes“ Metallnanopartikel angenommen. Dies ist insofern gerechtfertigt, da man in den gängigen Theorien die Spitze gerne als Dipol bzw. elongiertes

<sup>12</sup>engl: „scattering Scanning Nearfield Optical Microscopy“ (s-SNOM)



Partikel [58] betrachtet, und so vorerst der Einfluss des Cantilevers vernachlässigt wird, um so das System leichter verständlich zu machen. Das hier behandelte „s-SNOM“ besteht demnach aus einem mit der Frequenz  $\Omega_0$  frei schwingenden Partikel als „Spitze“ und einem Substrat mit dem Brechungsindex  $n = 1.5$  als Probe. In diesem Substrat ist ein weiteres Partikel als „Resonator“ eingebettet. Eine Skizze des Systems ist in Abbildung 5.2.12a dargestellt. Beide Partikel wurden mit DRUDE-LORENTZ-Gold (siehe Tabelle 2.3.1) mo-



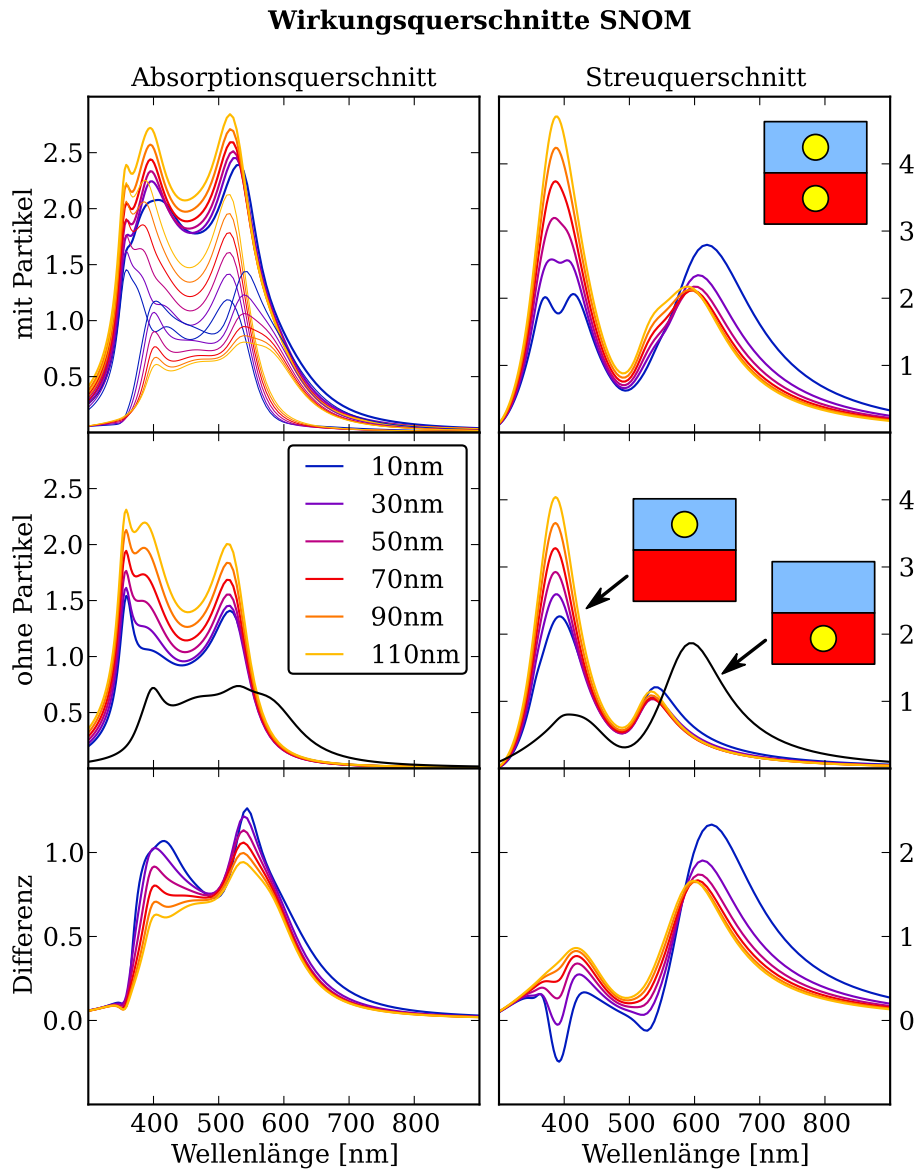
**Abbildung 5.2.12.:** 3D s-SNOM Schema. Das System ist in zwei Halbräume geteilt, der obere (blau) besteht aus Luft und der untere (rot) aus dem Substrat ( $n = 1.5$ ). Das obere Partikel schwingt mit einer Amplitude von 100 nm und approximiert die s-SNOM-Spitze, das untere Partikel dient als im Substrat eingebetteter Resonator.

dellert. Um die Schwingung der Spitze nachzubilden, wurde der Abstand zwischen dem Substrat und der Unterkante des oberen Partikels mit den Werten 10, 30, 50, 70, 90 und 110 nm abgetastet<sup>13</sup>. Dabei wurde das Streu- und Absorptionsspektrum im Bereich von 300 bis 900 nm in 120 Punkten mittels des DFT-Filters bestimmt. Als Anregung diente ein im Winkel von  $75^\circ$  einfallender p-polarisierter differenzierter GAUSSscher Puls mit einer Mittenwellenlänge von 450 nm. Neben dem System mit eingebettetem Partikel wurden als Referenzen die Rechnungen mit jeweils nur einem der beiden Partikel wiederholt, das

<sup>13</sup>Dabei wird angenommen, dass sich das elektromagnetische Feld instantan einstellt. Dies ist gerechtfertigt, da die Spitze typischerweise im kHz-Bereich schwingt, und das elektromagnetische Feld Antwortzeiten von wenigen fs besitzt.

heißt, einerseits als System mit schwingender „Spitze“ über dem gleichen Substrat, aber ohne eingebetteten Partikel und andererseits als System mit Substrat und eingebettetem Partikel, aber ohne schwingende „Spitze“. Bis auf das jeweils fehlende Partikel sind diese Systeme identisch mit Abbildung 5.2.12. Die so ermittelten Wirkungsquerschnitte sind in Abbildung 5.2.13 in Abhängigkeit von der Position der „Spitze“ und der Wellenlänge der Anregung zusammengefasst. An Hand der Spektren kann man für jedes Partikel zwei Streuresonanzen identifizieren, für die „Spitze“ liegen diese bei  $\approx 390$  und  $\approx 540$  nm und für das eingebettete Partikel bei  $\approx 400$  und  $\approx 600$  nm. Dabei spalten sich die „blauen“ Resonanzen bei Annäherung der „Spitze“ an das eingebettete Partikel auf, während sich die „rote“ Resonanz des eingebetteten Partikels um 40 nm in Richtung größerer Wellenlängen verschiebt und die „rote“ Resonanz der Spitze unterdrückt wird. Beides ist eine Folge der mit der Annäherung stärker werdenden Partikel-Partikel-Wechselwirkung. Die „blaue“ Resonanz des Referenzsystems (der „Spitze“) ist dabei keine Quadrupolresonanz, sondern ebenfalls ein Dipol, welcher durch eine Kombination des anregenden, des reflektierten Feldes und dem spektralen Verlauf des verwendeten Materialmodells hervorgerufen wird. Dies wird insbesondere durch die starke Abstandsabhängigkeit von der Oberfläche im Vergleich zur „roten“ Resonanz deutlich, wenn man beachtet, dass eine Quadrupolresonanz deutlich stärker an das Partikel gebunden ist, als eine Dipolresonanz. Dieser Eindruck kann zusätzlich durch die Feldbilder bestätigt werden (Hier nicht gezeigt.).

Beim Absorptionsquerschnitt ohne eingebettetes Partikel zeigen sich drei Maxima (bei 370, 400 und 520 nm), wobei das mittlere mit abnehmender Entfernung der „Spitze“ zum Substrat unterdrückt wird, diese ist wieder der „Interferenzdipol“. Die erste Resonanz (bei 370 nm) ist der Quadrupol und die letzte Resonanz (bei 520 nm) ist die Dipolresonanz. Betrachtet man nun das eingebettete Partikel ohne Spitze, so zeigt dieses eine Dipolresonanz in der Absorption bei ca. 550 nm und eine Quadrupolresonanz bei 400 nm. Dabei ist die Dipolresonanz sehr breit. Dies ändert sich, wenn man nun die Spitze annähert. So steigert man durch die Spitze generell die Absorption im unteren Partikel. Des Weiteren wird die Dipolresonanz der Absorption schmaler und stärker ausgeprägt. Jetzt lässt sich auch der Doppelpeak im Streuspektrum verstehen. So wird einerseits der Interferenzdipol unterdrückt, während die beiden Quadrupolresonanzen der beteiligten Partikel deutlicher sichtbar werden.



**Abbildung 5.2.13.:** Absorptions- und Streuquerschnitte des 3D s-SNOM. Oben: mit eingebettetem Partikel, Mitte: Referenz ohne eingebettetes Partikel und (in schwarz) nur das Substrat, Unten: die Differenz aus den Rechnungen mit eingebettetem Partikel abzüglich der Referenz ohne eingebettetes Partikel. Für das System mit „Spitze“ und eingebettetem Partikel ist der Absorptionsquerschnitt zusätzlich für beide Gold-Kugeln getrennt aufgetragen.

Um nun die harmonische Analyse der mit  $\Omega_0$  schwingenden „Spitze“ durchzuführen, müssen zuerst die distanzabhängigen Werte in zeitabhängige Werte umgerechnet werden. Dazu wird die Schwingungsgleichung, mit  $z_0$  als niedrigstem „Spitze“-Substrat-Abstand

und  $\hat{z}$  als Schwingungsamplitude, entsprechend umgeformt<sup>14</sup>:

$$z_j = z_0 + \frac{\hat{z}}{2} \cdot (1 - \cos(\Omega_0 t_j)) \quad (5.2.5a)$$

$$t_j = \frac{1}{\Omega_0} \arccos \left( 1 - 2 \frac{z_j - z_0}{\hat{z}} \right) \quad (5.2.5b)$$

Dabei wurde (willkürlich) angenommen, dass sich die „Spitze“ zu Beginn einer jeden Periode am unteren Umkehrpunkt befindet. Da nun die Messgrößen  $C$  als Funktion der Zeit vorliegen, kann eine Interpolation mittels der FOURIERreihe erfolgen:

$$C = \sum_k^n a_k \cos(k\omega_0 t) \quad (5.2.6)$$

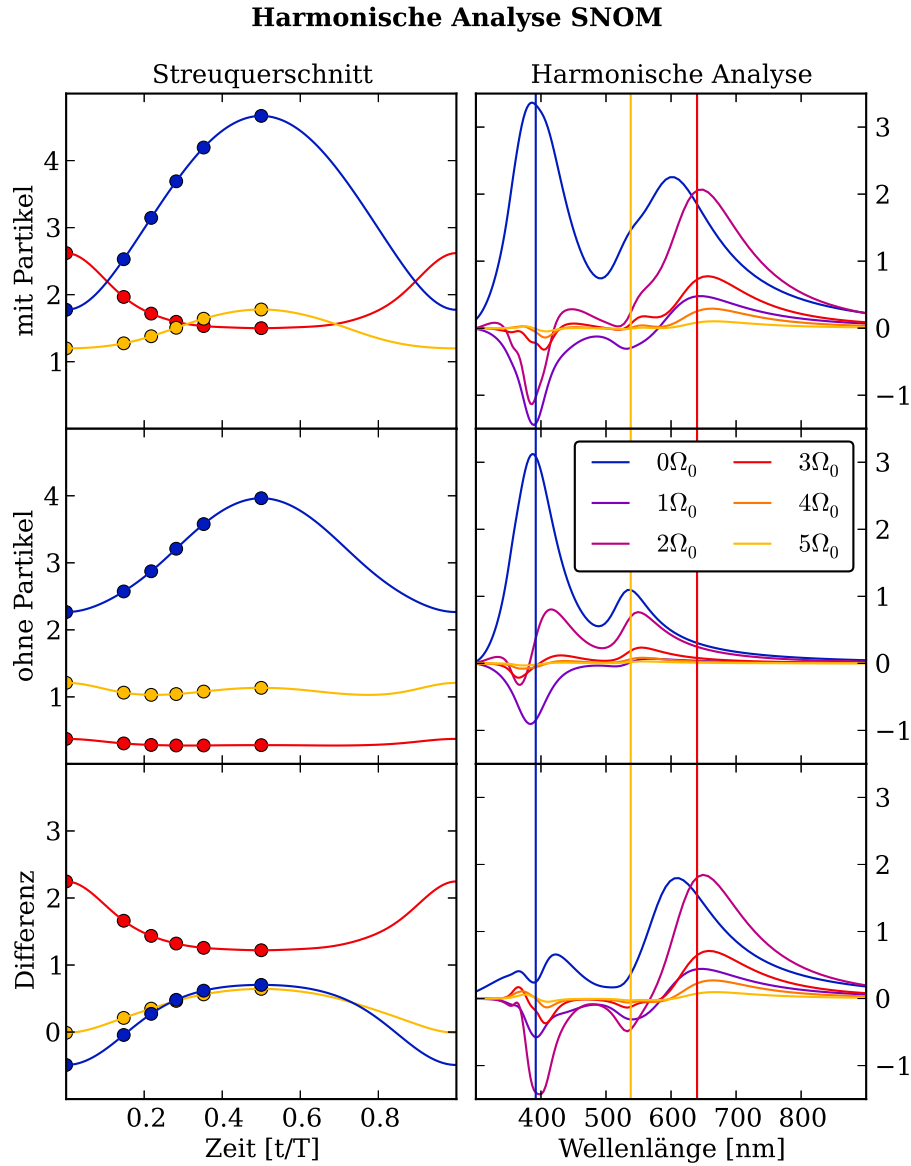
Dabei wurde gleichzeitig die Symmetrie der Messgrößen bezüglich der Umkehrpunkte berücksichtigt, das heißt, die Sinus-Terme in der FOURIERreihe entfallen. Mit  $C(z_j) = C(t_j) = C_j$  die Amplitude zur Zeit  $t_j$  ( $1 \leq j \leq m$ ) erhält man:

$$\sum_k^n a_k \cos(k\omega_0 t_j) = C_j \quad (5.2.7)$$

Wählt man nun die Ordnung der FOURIERreihe<sup>15</sup> genauso hoch wie die Anzahl der Substrat-„Spitze“-Distanzen ( $n + 1 = m = 6$ ), das heißt, man hat genauso viele Unbekannte wie Gleichungen, so erhält man ein quadratisches lineares Gleichungssystem für die Koeffizienten der FOURIERreihe. Dieses kann nun mit einem beliebigen Lösungsverfahren, zum Beispiel der GAUSSELIMINATION, gelöst werden. Die so ermittelte Zerlegung des Streuquerschnittes ist zusammen mit der FOURIERinterpolation als Funktion der Anregungswellenlänge in Abbildung 5.2.14 dargestellt. Vergleicht man nun die FOURIERzerlegung der Streuquerschnitte mit denen aus Abbildung 5.2.13 so wird die Bedeutung der einzelnen Frequenzkomponenten ( $a_k \hat{=} k\Omega_0$ ) deutlicher. So beschreibt  $0\Omega_0$  den Mittelwert des Streuspektrums und  $1\Omega_0$  die lineare Schwingungsamplitude des Streuspektrums. Dabei wird  $1\Omega_0$  negativ, wenn der Streuquerschnitt mit der Distanz zum Substrat abnimmt. Diese Änderung des Streuspektrums mit der Substrat-„Spitze“-Distanz ist dann besonders groß, wenn wenigstens eines der beteiligten Partikel an seiner Resonanz angeregt wird. Neben diesen sowohl Fern- als auch Nahfeldanteile enthaltenden FOURIERkoeffizienten kann man

<sup>14</sup>Man beachte, dass  $\arccos$  nur die erste Hälfte der Schwingung liefert, die zweite erhält man durch Spiegelung der Ergebnisse. Dies wird aber hier nicht benötigt, da im weiteren Verlauf die Symmetrie ausgenutzt werden kann.

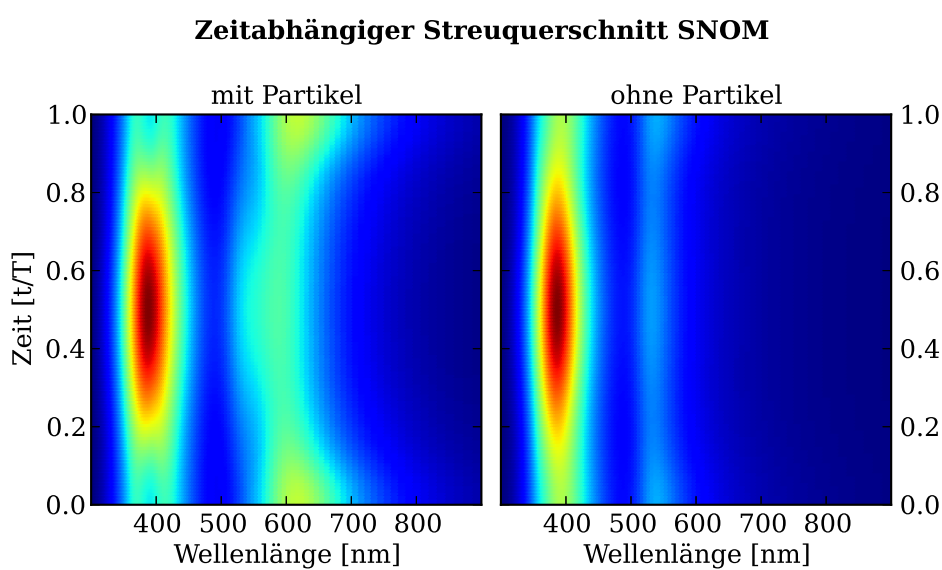
<sup>15</sup>Dabei muss beachtet werden, dass bei der FOURIERreihe analog zu den Basispolynomen die Nummerierung bei Null begonnen wird.



**Abbildung 5.2.14.:** Harmonische Analyse der Streuspektren. Links: die sechs zeitabhängigen Streuquerschnitte (Punkte) für die Anregungswellenlängen von 392 nm (blau), 538 nm (gelb) und 640 nm (rot) sowie die mit diesen Messpunkten ermittelten FOURIERinterpolationen (Linien). Rechts: die anregungswellenlängenabhängigen FOURIERkoeffizienten aus der harmonischen Analyse der Partikelschwingung. Die nichtlinearen Koeffizienten ( $> 1\Omega_0$ ) sind um den Faktor zehn überhöht dargestellt. Zusätzlich sind noch die Wellenlängen der links gezeigten zeitabhängigen Streuquerschnitte markiert.

auch die höheren Harmonischen betrachten. Diese sind speziell dort stärker ausgeprägt, wo das System seine Streuresonanzen besitzt, das heißt, im kombinierten System bei 392 und

640 nm und im Referenzsystem (ohne eingebettetem Partikel) hauptsächlich bei 538 nm. Nun könnte man sich fragen, wieso die Streuresonanz des Referenzsystems, welche sich bei  $\approx 390$  nm befindet, und sich stark mit dem Abstand von der Oberfläche ändert, nur einen, im Vergleich zur ersten Harmonischen, schwach ausgeprägten Anteil in den höheren Harmonischen zeigt. Die Ursache liegt darin begründet, dass es sich bei dieser Resonanz, wie oben bereits erläutert, nicht um eine „klassische“ Nahfeldresonanz des Partikels, sondern um ein durch Interferenz und Streuung hervorgerufenen Fernfeldeffekt handelt. Neben der harmonischen Analyse der zeitabhängigen Streuspektren lohnt es sich ebenfalls die FOURIERinterpolation direkt anzuschauen. Diese ist für ausgewählte Wellenlängen (392, 528 und 640 nm) in Abbildung 5.2.14 zusammen mit den ursprünglichen Messpunkten aufgetragen. Auffällig ist der im Zeitverlauf nicht äquidistante Abstand der Messwerte als Folge der Ort-zu-Zeit-Transformation nach Gleichung 5.2.5. Des Weiteren lässt sich hier gut die Symmetrie erkennen, welche bei der FOURIERzerlegung (Gleichung 5.2.6) ausgenutzt wurde. Deutlich ist der nichtsinusförmige Verlauf der Streuung bei 640 nm Anregungswellenlänge für das System mit eingebettetem Partikel zu sehen (vergleiche dazu auch Abbildung 5.2.11). Auch ist bei 528 nm die Oberschwingung für das System ohne eingebettetem Partikel sichtbar. Beide deuten auf den Nahfeldanteil der Dipolresonanzen der Partikel hin. Dem gegenüber sieht die rote Resonanz (bei 640 nm) deutlich sinusförmiger aus, was sich auch in der harmonischen Analyse widerspiegelt. Insbesondere zeigt der „Interferenzdipol“ des Systems ohne eingebettetem Partikel nur eine schwache Anregung der höheren Harmonischen. Diese für ausgewählte Wellenlängen gezeigte FOURIERinterpolation lässt sich auch auf das gesamte Spektrum erweitern. Es ist in Abbildung 5.2.15 dargestellt. Deutlich lässt sich hier die Verschiebung der Streuresonanz des eingebetten Partikels (bei ca. 640 nm) über den Partikelabstand beziehungsweise Zeitverlaufes nachvollziehen. Auch kann man feststellen, dass sich die Aufspaltung der Resonanz bei (ca. 390 nm) nur in einem kurzen Zeitbereich vollzieht. Dies verdeutlicht noch einmal den starken anharmonischen Anteil und damit die Bedeutung des Nahfeldes für die Partikel-Partikel-Wechselwirkung.



**Abbildung 5.2.15.:** Rekonstruierte zeitabhängige Streuspektren des 3D s-SNOM. Deutlich lässt sich die späte Aufspaltung des Streusignals und die Resonanzverschiebung bei dem System mit Partikel im Vergleich zu dem System ohne Partikel erkennen.

### 5.3. 3D - SHG - Spektren

Ogleich die linearen Rechnungen bereits einen weitgefaßten Überblick über die Möglichkeiten der hier verwendeten Methode geben, so war die zeitabhängige Behandlung der MAXWELL-Gleichungen bisher nicht erforderlich. Dies ändert sich nun, da zu nichtlinearen Eigenschaften plasmonischer Systeme übergegangen wird. Speziell handelt es sich um die Erzeugung der zweiten Harmonischen (SHG, siehe Abschnitt 2.5).

#### Vorbemerkungen

Analog zur MIE-Theorie [42], welche die Absorptions- und Streuspektren kolloidaler Partikel erklären konnte, stellt sich nun die Frage nach dem spektralen Verlauf der SHG. Dabei wird schnell klar, dass breitbandige, das heißt, mit einem sehr kurzen Puls gewonnene, Spektren nicht zum Ziel führen, da die so erhaltenen Ergebnisse stets eine Kombination der linearen und deutlich schwächeren nichtlinearen Anteile beinhalten. Typischerweise lässt sich der SHG-Beitrag kaum vom viel höheren linearen Hintergrund isolieren.

Ein Ausweg besteht darin, spektral schmale Pulse mit variabler Mittenwellenlänge  $\lambda$  zu verwenden. Deren Amplitude bei der halben Wellenlänge  $0.5\lambda$  ist praktisch Null, so dass der lineare Streuquerschnitt bei der SHG-Frequenz keinen Beitrag mehr liefert. Tastet man mit diesem Puls nun das gesamte Spektrum ab, so erhält man für jede Rechnung die Absorptions- und Streuspektren bei  $\lambda$  und die SHG-Streuspektren bei  $0.5\lambda$ . Der Nachteil dieses Verfahrens besteht darin, dass jeder Frequenzpunkt, analog zu einer Frequenzdomänenmethode, einzeln berechnet werden muss und nicht mehr wie in den linearen Rechnungen zuvor alle Frequenzpunkte durch Filterung aus einer einzigen Simulation gewonnen werden können.

Bei den nachfolgenden Rechnungen wurde ein  $1/\Delta\omega = 13.3$  fs langer GAUSSscher Puls (siehe Abschnitt 2.6.3) mit einer maximalen Anregungsfeldstärke von  $10^9 \frac{V}{m}$  und einer Trägerwellenlänge  $\lambda$  zwischen 200 und 800 nm in Schritten von 5 nm verwendet. Als Materialmodell für die plasmonischen Systeme wurde das in Abschnitt 2.3.3 vorgestellte Hydrodynamische Modell verwendet. Dabei wurden für die Ermittlung der Elektronendichte und der Dämpfung auf die Werte des DRUDE-Silbers nach Tabelle 2.3.1 zurückgegriffen. Die effektive Masse der Elektronen wurde mit  $1.111253 \cdot m_e$  angenommen.



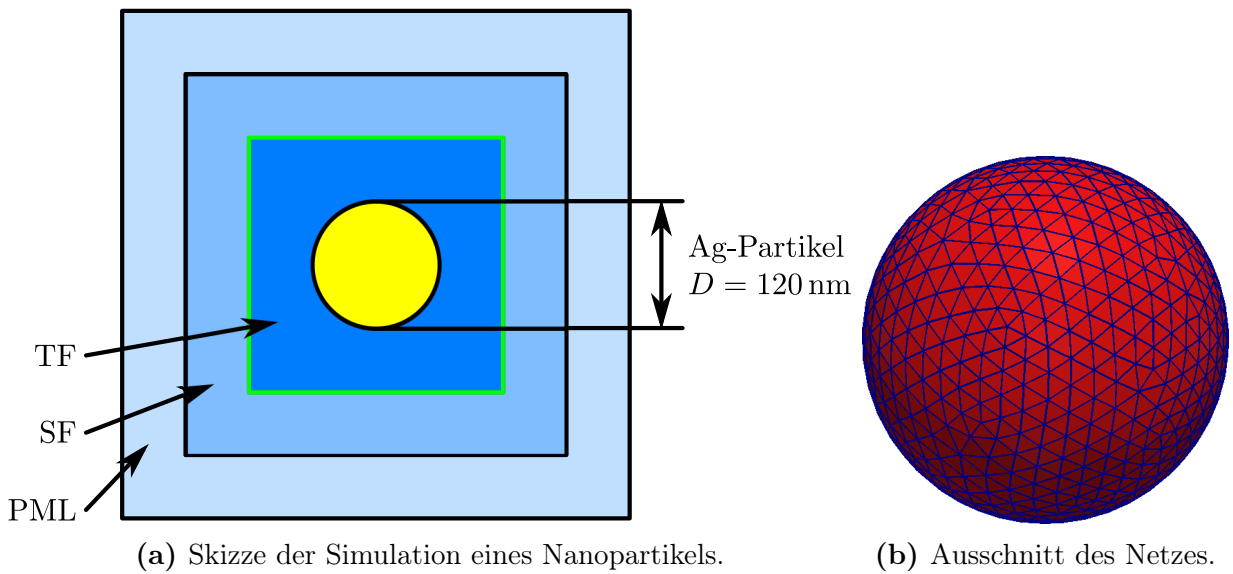
Die nachfolgenden Rechnungen entstanden in Kooperation mit der Arbeitsgruppe um Prof. Dr. BUSCH am KIT in Karlsruhe. Dabei wurde Code verwendet, welcher maßgeblich, aber nicht ausschließlich, von Jens Niegemann, Michael König und Christopher Prohm entwickelt wurde. Durch die große Anzahl an Rechnungen, 122 Simulationen je Spektrum, wurden insgesamt 10 CPU-Jahre Rechenzeit aufgewendet. Hier wurde allerdings nicht mit runden Elementen gerechnet. Um diesen Nachteil auszugleichen, wurden die Netze relativ fein mit  $h_{min} = r/8$  diskretisiert. Kritisch ist hier ohnehin nicht die Plasmonwellenlänge, sondern die FERMIwellenlänge der Elektronen, welche sich im sub-Nanometer-Bereich befindet.

Im Folgenden werden zwei Systeme behandelt, einerseits das sphärische Nanopartikel, andererseits ein System aus zwei Nanopartikeln mit variablem Abstand. Dabei wird gezeigt, dass sich die SHG-Effizienz teilweise mit Hilfe der linearen Wirkungsspektren vorhersagen lässt.

### 5.3.1. Einzelnes Partikel

Eines der einfachsten plasmonischen Systeme stellt das Einzelnanopartikel da. Darüber hinaus sind die linearen Eigenschaften in den vorangegangenen Kapiteln recht ausführlich behandelt worden. Daher soll dieses System auch bei den nichtlinearen Rechnungen am Anfang stehen. Für das Material des Nanopartikels wird, wie oben beschrieben, Silber angesetzt. Der Durchmesser wurde mit 120 nm festgelegt. Der prinzipielle Aufbau der Rechnung ist analog zu den vorangegangenen in Abschnitt 5.2.1. Eine Skizze des Systems wird zusammen mit einem Ausschnitt des Netzes in Abbildung 5.3.1 gezeigt. Die Wirkungsspektren sind in Abbildung 5.3.2 zusammengestellt. Zu beachten ist, dass jedes Maximum (jede Farbe) für eine Simulation, also einen Frequenzpunkt steht. Die Einhüllenden dieser einzelnen Ergebnisse führen, im linearen Fall, wieder auf die gewöhnlichen Streu- und Absorptionsspektren. Hier wurden nun die einzelnen Maxima für die Absorption und die Streuung von Eins (Dipolresonanz) bis Fünf (Dekapol) durchnummeriert. Dies erleichtert später die Zuordnung der verschiedenen linearen zu den nichtlinearen Resonanzen. Da die linearen Eigenschaften denen aus den vorangegangenen Abschnitten 5.2.1 ähneln, soll hier nicht noch einmal detailliert darauf eingegangen werden. Es ist jedoch festzustellen, dass trotz des Wechsels vom DRUDE- zum Hydrodynamischen Materialmodel die linearen Eigenschaften weitestgehend, das heißt, qualitativ<sup>16</sup>, erhalten bleiben. Die Voraussetzung

<sup>16</sup>Bei den hier gezeigten Rechnungen kam es gegenüber dem reinen DRUDE-Model zu einer kleinen Verschiebung der Resonanzen und erhöhten Absorptionswerten, deren Ursache in den nichtlokalen Effekten des Elektronenfluids vermutet wird. Die Erforschung der Ursachen ist noch nicht abgeschlossen.

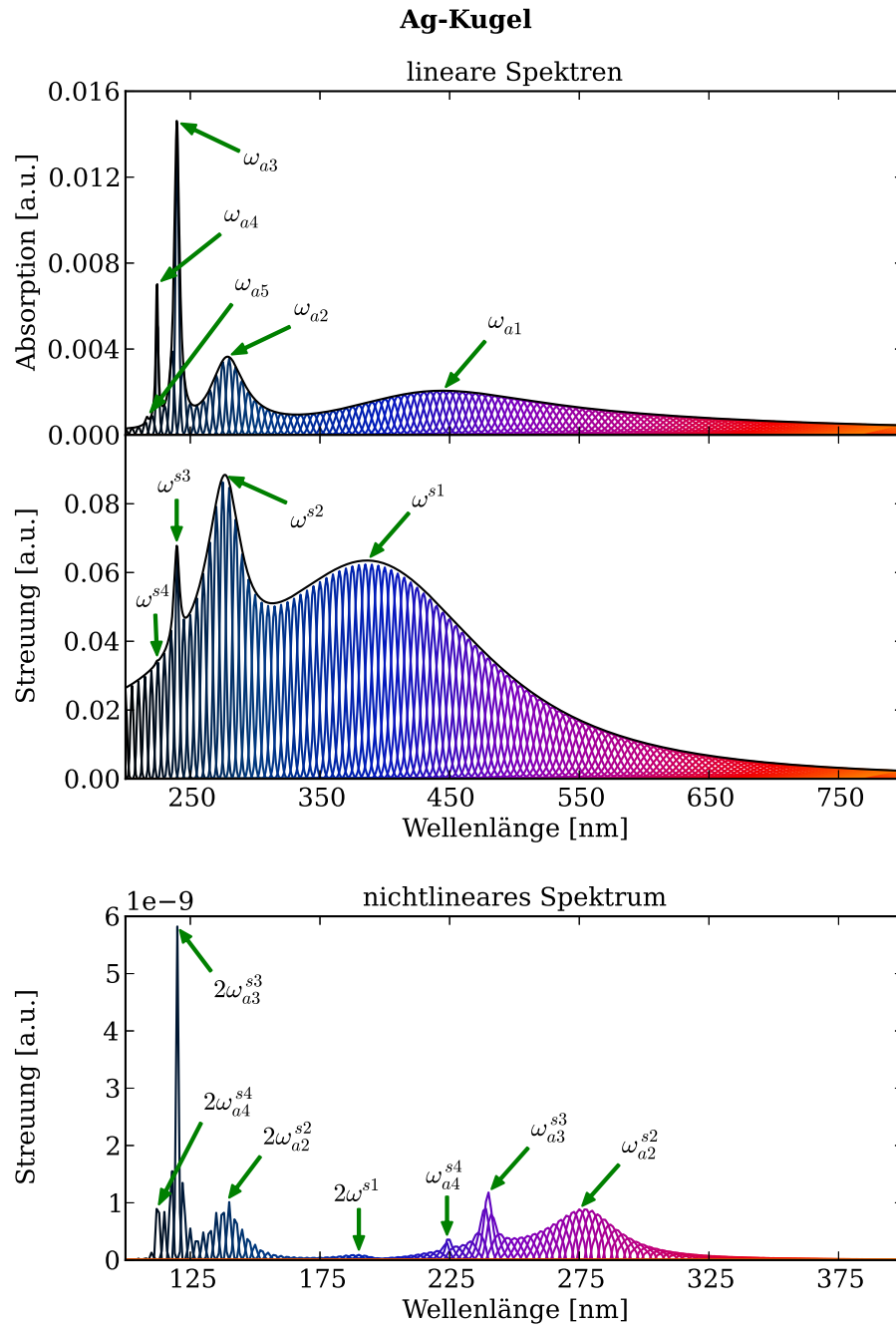


**Abbildung 5.3.1.:** Skizze der Simulation eines Nanopartikels und Ausschnitt des verwendeten Netzes.

ist natürlich, dass die nichtlinearen Effekte einen vernachlässigbaren Einfluss auf die linearen Ergebnisse besitzen. Dies ist hier insofern gegeben, da die nichtlinearen Streuwerte ca. sieben Größenordnungen unter den linearen liegen.

Betrachtet man das nichtlineare Spektrum, so fällt zuerst auf, dass es ebenso wie das lineare verschiedene Maxima aufweist. Erstaunlicherweise lässt sich dies nun recht präzise den linearen Resonanzen zuordnen. Dabei wird noch unterschieden, ob das Maxima auftritt, wenn das Partikel bei der fundamentalen Anregungswellenlänge oder bei der SHG-Wellenlänge eine lineare Resonanz aufweist. Erstere sind mit  $2\omega$ , letztere mit  $\omega$  und der Nummer der betreffenden Resonanz im Spektrum gekennzeichnet. Dabei verwundert es weniger, dass die SHG besonders ausgeprägt ist, wenn das Partikel bei der Anregungsfrequenz resonant ist. Denn hierbei ist das Feld am und im Partikel besonders hoch, was sich vorteilhaft auf die, quadratisch von der Feldstärke abhängende, SHG-Effizienz auswirkt. Erstaunlicher ist jedoch, dass die SHG auch dann erhöht sein kann, wenn die Wellenlänge des Anregungslichts fern einer linearen Resonanz ist und stattdessen die erzeugte SHG eine Partikelresonanz trifft. Dabei hat es den Anschein, dass dieser Effekt nur bei höheren Resonanzen jenseits des Dipols auftritt.

Sollte sich dieses Verhalten bei weiteren Systemen bestätigen, könnte man so von den linearen Eigenschaften auch auf die nichtlinearen Eigenschaften eines Partikels schließen, was die experimentelle Behandlung deutlich vereinfachen würde.

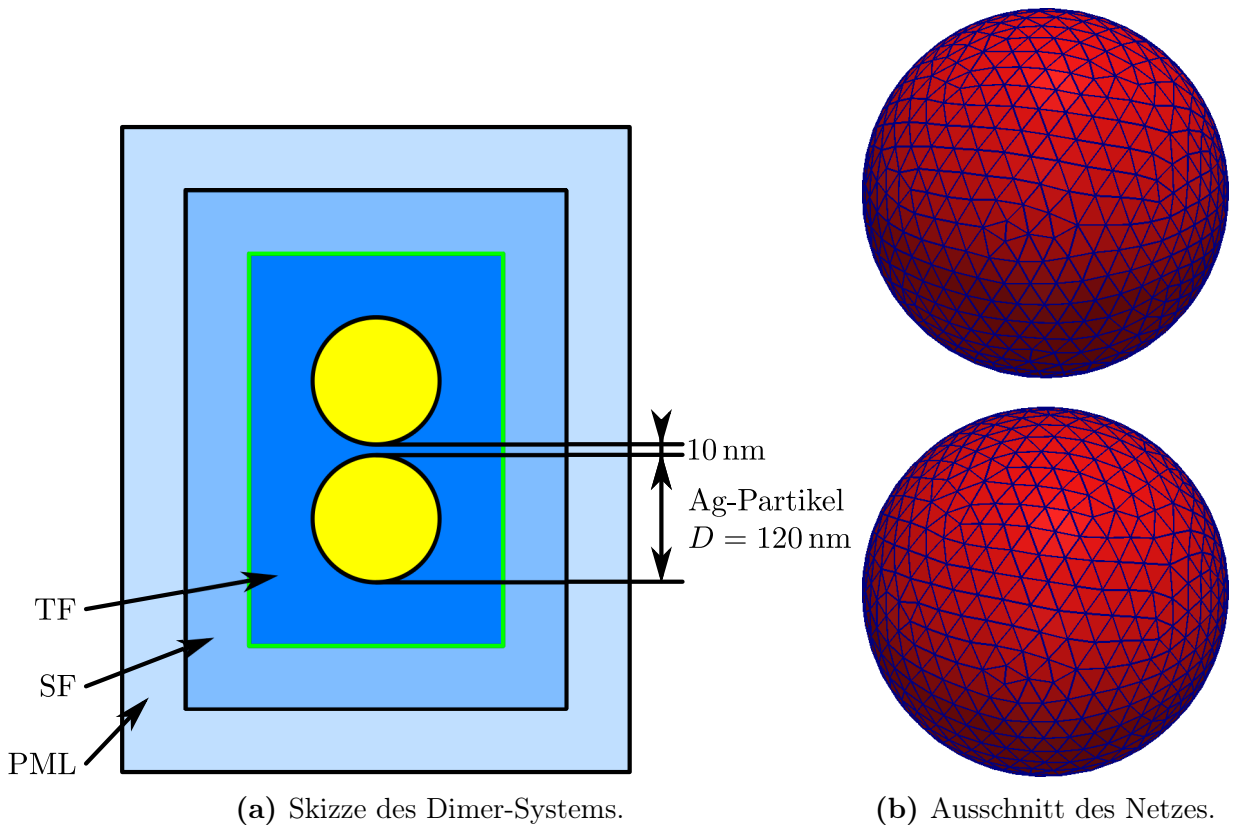


**Abbildung 5.3.2.:** Streu- und Absorptionsspektren der linearen und nichtlinearen Partikelantwort. Jede Farbe steht für eine Simulation. Die Einhüllende (schwarz) der linearen Spektren ergibt wieder die „bekannten“ linearen Streu- und Absorptionsspektren. Die einzelnen Resonanzen wurden der Reihe nach durchnummeriert.

### 5.3.2. Dimere

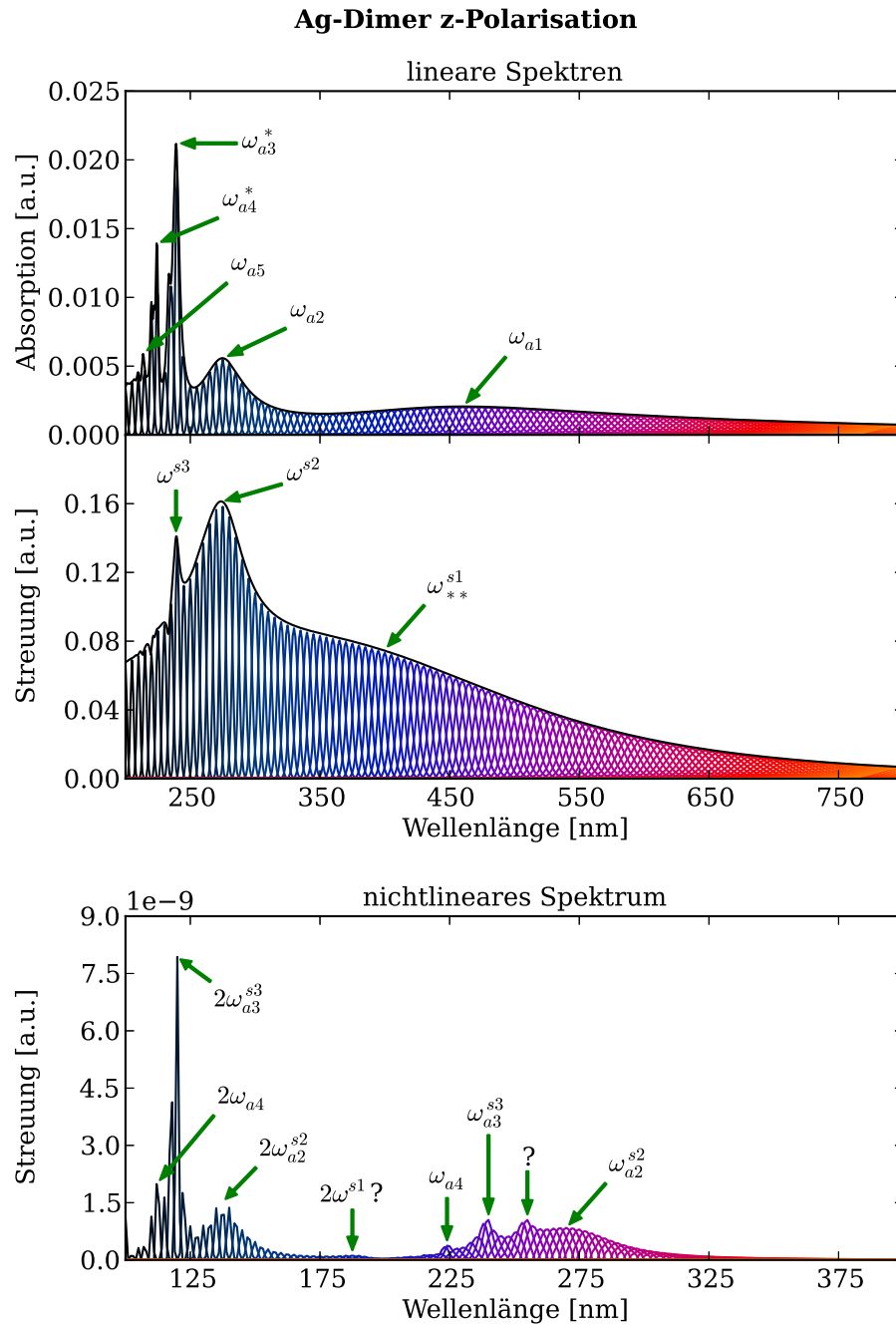
Um die im letzten Abschnitt aufgestellte Vermutung zu überprüfen, wird als nächstes ein System mit zwei Partikeln behandelt. Dabei hat man den Vorteil, dass sich die linearen Resonanzen in Abhängigkeit vom Partikelabstand und der Polarisation des einfallenden Lichtes einstellen lassen, und man so eine Verschiebung des nichtlinearen Spektrums provozieren kann. Darüber hinaus kann man versuchen, das System so einzustellen, dass es sowohl bei der Anregungswellenlänge als auch bei der SHG-Wellenlänge eine Resonanz besitzt.

Das hier verwendete Dimer besteht aus zwei identischen Silberpartikeln mit gleichen Durchmessern von 120 nm und einem Abstand von 10 nm. Dabei wurden zwei Polarisationen gerechnet, einerseits entlang der Verbindungsachse, der „langen“ Achse, der beiden Partikel, hier „y“-Polarisation genannt, und andererseits senkrecht zu dieser, entlang der „kurzen“ Achse, hier „z“-Polarisation genannt. Eine Skizze der Versuchsanordnung und eines Netzausschnittes findet sich in Abbildung 5.3.3. Begonnen wird nun mit der Polari-



**Abbildung 5.3.3.:** Skizze des Dimer-Systems und Ausschnitt des Netzes.

sation entlang der „kurzen“ Achse. Die Spektren sind in Abbildung 5.3.4 dargestellt. Im Gegensatz zum Einzelpartikel ist die Dipolresonanz  $\omega^{s1}$  im Streuspektrum leicht unter-

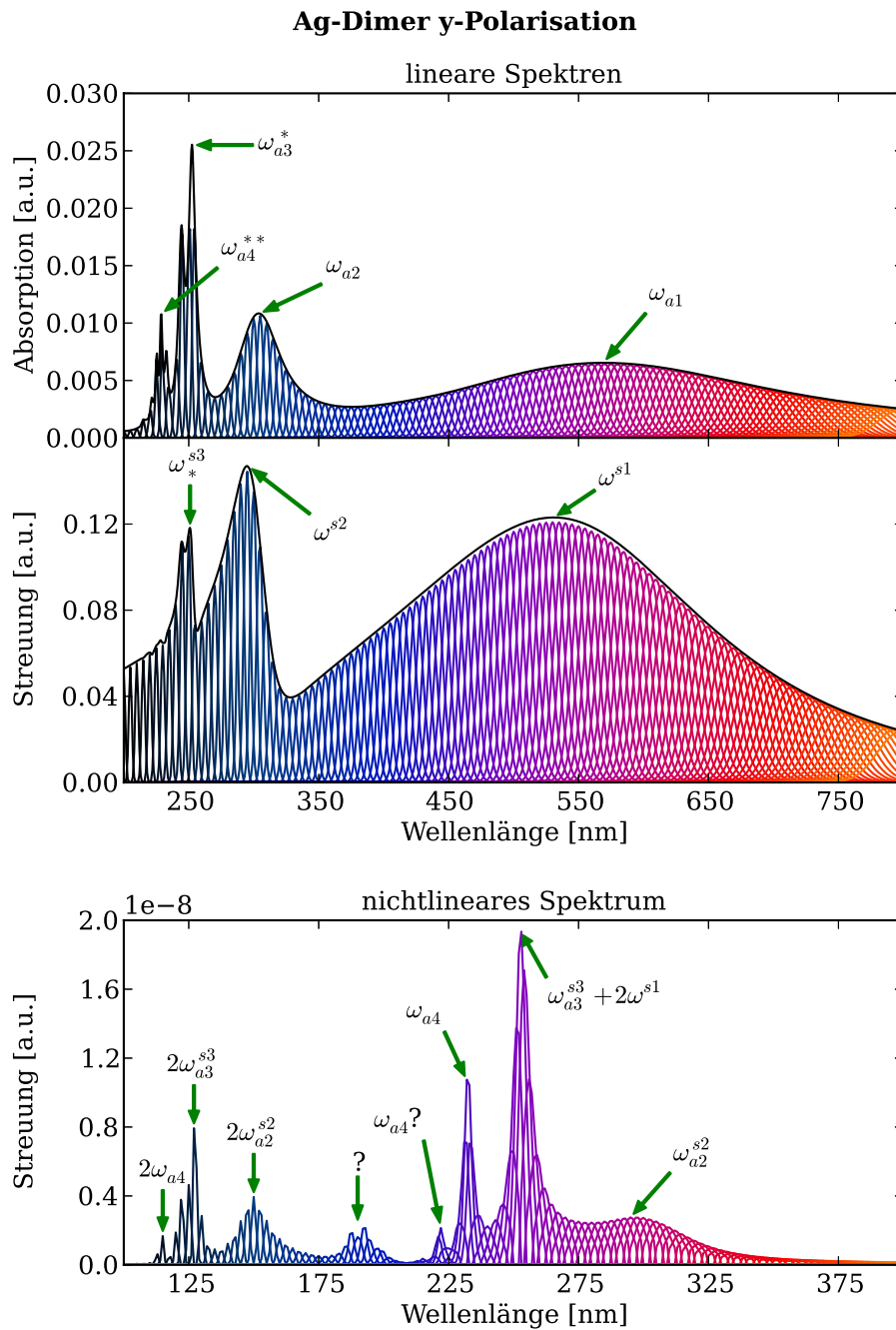


**Abbildung 5.3.4.:** Streu- und Absorptionsspektren der linearen und nichtlinearen Partikelantwort. Jede Farbe steht für eine Simulation. Die Einhüllende (schwarz) der linearen Spektren ergibt wieder die „bekannten“ linearen Streu- und Absorptionsspektren. Die einzelnen Resonanzen wurden der Reihe nach durchnummeriert. (\* Doppelpeak, \*\* Schulter, „?“ nicht/schlecht zuzuordnen)

drückt, was sich in der Schulter bemerkbar macht. Darüber hinaus bilden sich unterhalb von  $\lambda < 250 \text{ nm}$  im Absorptionsspektrum Doppelpeaks aus. Analog zum einzelnen Nanopartikel zeigt auch hier die SHG eine erhöhte Abstrahlung, wenn die Anregungswellenlänge auf eine lineare Resonanz trifft. Im Gegensatz zum Einzelpartikel ist hier der Dipol  $\omega_{**}^{s1}$  so schwach ausgeprägt, dass auch das zugehörige SHG-Signal  $2\omega^{s1}$  nur sehr gering ausgebildet ist. Betrachtet man nun den Fall, dass die SHG-Wellenlänge auf einer linearen Resonanz liegt, so stellt man erneut fest, dass es auch hier zu einer Verstärkung des SHG-Signals kommen kann. Zusätzlich taucht jetzt ein Maximum auf, das sich keiner linearen Resonanz zuordnen lässt (markiert mit „?“). Dies ist gleichzeitig auch das Maximum, welches bei dem einzelnen Nanopartikel, im Gegensatz zu den drei anderen, nicht auftaucht.

Abschließend folgt nun das Dimer mit „y“-Polarisation (siehe Abbildung 5.3.5). Hier ist die lineare Resonanz verglichen mit dem Einzelpartikel zu längeren Wellenlängen hin verschoben. Am deutlichsten ist dies bei der Dipolresonanz  $\omega^{s1}$ , bei der die Verschiebung ca. 175 nm beträgt, zu erkennen. Auch hier lassen sich nun fast alle nichtlinearen Maxima den linearen Resonanzen  $\omega$ , beziehungsweise deren verdoppelter Frequenz  $2\omega$ , zuordnen. Insbesondere ist das SHG-Maximum bei 255 nm, bei dem das lineare System sowohl bei der Anregungswellenlänge (510 nm) wie auch der SHG-Wellenlänge resonant ist, besonders stark ausgeprägt. Genauso wie bei der „z“-Polarisation taucht auch hier ein Maximum auf, welches sich keiner linearen Resonanz zuordnen lässt. Um diesem Sachverhalt auf den Grund zu gehen, werden nun beide Polarisationen nebeneinander betrachtet. Dabei fällt auf, dass die nicht zuordenbaren SHG-Maxima, dem linearen Streudipol beziehungsweise dessen SHG-Wellenlänge, der jeweils anderen Anregung entsprechen. Dies legt die Vermutung nahe, dass die SHG bei allen, und nicht nur bei den zur Anregung gehörenden, linearen Resonanzen, eine erhöhte Effizienz zeigt. Ein ähnlicher Effekt ist bei der SHG an Metallfilmen in [8] und [5] dokumentiert.

In diesen ersten Voruntersuchungen zur wellenlängenabhängigen SHG wurde gezeigt, dass mit Hilfe der vollständigen linearen Absorptions- und Streuspektren Rückschlüsse auf die SHG-Effizienz gezogen werden können. Dies führt gerade in der Praxis zu einer deutlichen Vereinfachung, da lineare Spektren leichter messbar sind als nichtlineare Spektren. Auch gibt es deutliche Hinweise darauf, dass man die SHG plasmonischer Systeme durch Einstellen ihrer linearen Resonanzen optimieren kann. Insbesondere lässt sich die SHG-Ausbeute durch eine Kombination aus linearen Resonanzen bei Anregungs- und SHG-Wellenlänge erhöhen.



**Abbildung 5.3.5.:** Streu- und Absorptionsspektren der linearen und nichtlinearen Partikelantwort. Jede Farbe steht für eine Simulation. Die Einhüllende (schwarz) der linearen Spektren ergibt wieder die „bekannten“ linearen Streu- und Absorptionsspektren. Die einzelnen Resonanzen wurden der Reihe nach durchnummeriert. (\* Doppelpeak, \*\* Dreifachpeak, „?“ nicht/schlecht zuzuordnen)





## 6. Zusammenfassung

### 6.1. Diskussion

In dieser Arbeit wurde ein DG-Verfahren zur Simulation nanooptischer Systeme erstellt. Dabei wurde ein besonderer Wert auf die oberflächenkonforme Diskretisierung des Simulationsgebietes gelegt. Es hat sich gezeigt, dass, speziell im Bereich metallischer Nanostrukturen, dieses Vorgehen zu einer deutlichen Steigerung der Genauigkeit führt. Gleichzeitig konnten damit unnatürlich hohe Feldverstärkungen vermieden werden. Dies ermöglicht eine vertrauenswürdigere Vorhersage von Nahfeldüberhöhung und Molekülabsorption im Nahbereich von plasmonischen Strukturen. Darüber hinaus konnte bei den hier untersuchten Partikelsystemen die Rechenzeiten um den Faktor  $> 2$  verkürzt werden. Nachteilig bei dem hier angewendeten Verfahren ist die damit verbundene gestiegene Anforderung an die Qualität des Netzes, welche einen signifikanten Einfluss auf die Rechenzeit zeigte.

Als zweiter großer Themenkomplex wurde das spektrale Verhalten der SHG untersucht. Dabei wurde gezeigt, dass sich mit Hilfe der linearen Streu- und Absorptionsspektren die Effizienz der SHG gut vorhersagen lässt. Dabei kristallisieren sich zwei Fälle heraus. Zum einen derjenige, bei dem die Anregungsfrequenz eine lineare Resonanz trifft. Zum anderen, weitaus interessanter, derjenige, bei dem die erzeugte SHG auf einer solchen Resonanz liegt. Für die Vorhersage ist es wichtig, alle linearen Resonanzen des Partikelsystems gut zu kennen, da, bedingt durch die Anregung, einige unterdrückt werden können. Zusätzlich lässt sich die SHG-Effizienz weiter steigern, wenn man das Partikelsystem so einstellt, dass sowohl die Fundamentale, wie auch die SHG-Wellenlänge auf einer linearen Partikelresonanz liegen.

## 6.2. Ausblick

Für die weiteren Untersuchungen müssten die bisher verfügbaren Netzgeneratoren besser auf die Anforderungen einer Zeitschrittmethode hin optimiert werden, denn es hat sich herausgestellt, dass die Güte des Netzes den signifikantesten Einfluss auf die Rechenzeit besitzt. Dies ist umso bedeutender, wenn mit runden Elementen gerechnet werden soll.

Für die Simulation der SHG wäre es wünschenswert, die Oberfläche viel feiner diskretisieren zu können, um in den Bereich der Fermiwellenlänge der Elektronen auflösen zu können. Dies wird nur gelingen, wenn man das DG-Verfahren hochparallelisiert. Ansätze dazu gibt es bereits. Der nächste Schritt wäre, sich vom Potentialtopfmodell zu lösen und diesen durch einen realistischeren Potentialverlauf zu ersetzen. Dadurch sollte es möglich sein, ergänzende Aussagen über die Elektronendynamik am Rand eines Nanopartikels treffen zu können.

Da die ersten numerischen „Experimente“ zur Vorhersage der SHG-Effizienz recht erfolgreich verlaufen sind, ist es an der Zeit, diese durch das reale Experiment bestätigen zu lassen. Hier kann man sich zu nutze machen, dass sich die linearen Resonanzen in plasmonischen Systemen gut einstellen lassen. Auch steht sofort die Frage im Raum, wie eine Struktur beschaffen sein müsste, die sowohl bei der fundamentalen, wie auch der frequenzverdoppelten Wellenlänge ausgeprägte Resonanzen besitzt.

# A. Anhang

## A.1. Erzeugung der Raumdiskretisierung

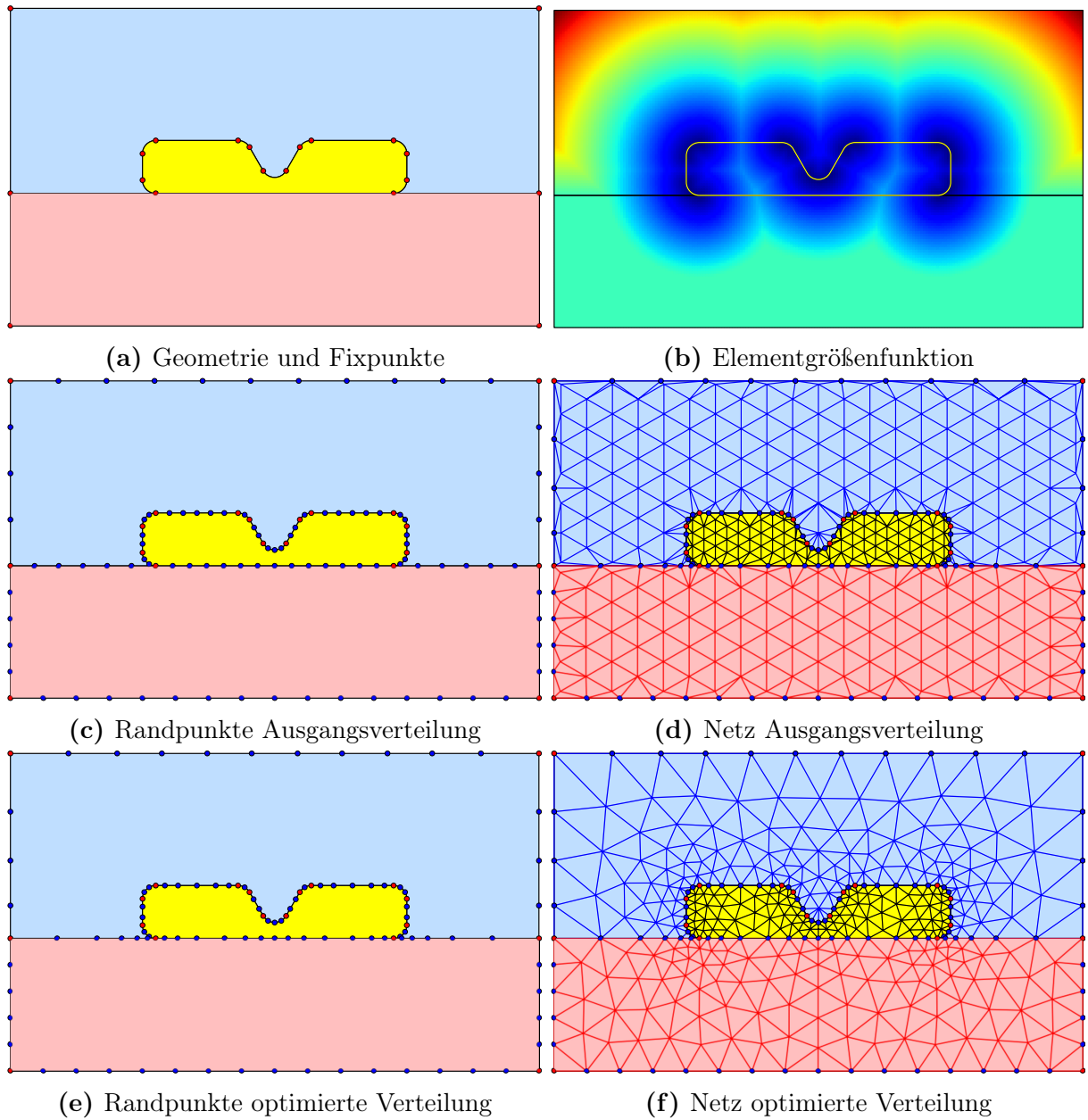
In diesem Abschnitt wird die Erzeugung der räumlichen Diskretisierung im 2D behandelt. Dabei kam ein selbstentwickelter Netzgenerator zum Einsatz. Seine Funktionsweise basiert auf dem „Distmesh“-Algorithmus aus [54], und wurde für die Zwecke dieser Arbeit modifiziert. Im Folgenden werden die einzelnen Stufen der Netzerzeugung, welche in Abbildung A.1.1 gezeigt sind, diskutiert. Begonnen wird mit einer Beschreibung der Geometrie, dabei kommen Volumenkörper (im 2D: Kreis, Rechteck, Polygon) zum Einsatz, die mittels boolescher Operationen verknüpft werden (siehe Abbildung A.1.1a). Zusätzlich benötigt man noch Diskretisierungsparameter wie:

- Größe der Elemente im Vakuum
- Anzahl der Elemente je Radius
- Stärke des Netzgradienten
- Lokale Materialparameter ( $\epsilon_r$ ,  $\mu_r$ )

Mit Hilfe dieser lässt sich eine Funktion konstruieren, die für jeden Raumpunkt die charakteristische Größe der Elemente bestimmt. Dabei werden zuerst die Elementgrößen an den gekrümmten Kanten festgelegt. Danach werden die lokalen, von Materialparametern abhängigen, Elementgrößen bestimmt<sup>1</sup>, als nächstes werden die unterschiedlichen Elementgrößen mittels des Netzgradienten angeglichen, um einen weichen Größenverlauf zu gewinnen. Der Kehrwert dieser Elementgrößenfunktion lässt sich als Element- beziehungsweise Punktdichte interpretieren. Für den Fall eines V-Grabens (siehe Abschnitt 5.1.2) auf einem Substrat ist die Funktion in Abbildung A.1.1b gezeigt. Integriert man die Elementdichte über eine Kante, erhält man die Anzahl der notwendigen Punkte für eine Kante. Dies ist für eine initiale äquidistante Punktverteilung in Abbildung A.1.1c dargestellt. Mit dem in

---

<sup>1</sup>Dies ist notwendig, da in Materialien mit einem höheren Brechungsindex die Wellenlänge kürzer ist, und damit die Auflösung der Felder beziehungsweise des Netzes verfeinert werden muss.



**Abbildung A.1.1.:** Schritte der Netzerzeugung.

Abschnitt A.1.1 beschriebenen Optimierungsverfahren, wird diese Punktverteilung nun so verschoben, dass die Elementdichte der Vorgabe entspricht (A.1.1e). Schließlich wird das Volumen in Dreiecke unterteilt<sup>2</sup>. Dabei wird zuerst eine initial Verticeverteilung vorgegeben. Hier wurde eine regelmäßige Verteilung gewählt<sup>3</sup> (siehe A.1.1d). Abschließend wird

<sup>2</sup>Dabei nutzt man die DELAUNAY-Triangulation. Bei diesem wird eine Triangulation so erzeugt, dass im Umkreis eines jeden Elementes kein weiterer Vertex liegt. Damit nimmt gleichzeitig der kleinste Elementinnenwinkel den größten (mittels dieser Punkteverteilung möglichen) Wert an.

<sup>3</sup>Andere Verfahren nutzen bereits bei der initialen Verticeverteilung die Elementgrößenfunktion. In [54] wurde dabei mit einem Zufallszahlen basierten Filter gearbeitet, diese haben dann Vorteile wenn es

auch diese Vertexverteilung mit dem nachfolgend (Abschnitt A.1.1) beschriebenen Verfahren optimiert. Das Resultat ist in Abbildung A.1.1f gezeigt. Zum Schluss können, wie in Kapitel 3 und 4 gezeigt, die Knoten in den Elementen bestimmt werden.

Das hier vorgestellte Verfahren verhindert keine spitzen Tangentenwinkel an gekrümmten Kanten. Eine Möglichkeit, dies zu berücksichtigen, ist die Krümmung der Elemente in das Volumen mit hineinzuziehen. Ein solches Verfahren ist in [53] geschildert.

### A.1.1. Optimierungsverfahren

Bei diesem Optimierungsverfahren nutzt man Ansätze aus der Mechanik. Dabei werden die Kanten zwischen den Punkten als Federn interpretiert. Dazu wird die Federkraft um so größer gewählt, je größer die Elementgröße an der Position der Kante sein soll<sup>4</sup>. Mit  $l$  der aktuellen Kantenlänge,  $\rho$  der gewünschten Kantenlänge und der „Federkonstante“  $k$  wird zunächst die Größe der entspannten Feder  $l_0$  berechnet:

$$l_0 = \rho * k * \sqrt{\frac{\sum l_i^2}{\sum \rho_i^2}}$$

Diese Grundlänge gleicht dabei den Unterschied zwischen den gewünschten und den tatsächlich vorhandenen Kantenlängen aus. Mit Hilfe dieser Grundlänge lautet die Federkraft  $F$  dann:

$$F = \min((l - l_0), l)$$

Dabei stellt das „min“ sicher, dass die Kraft ausschließlich repulsiven Charakter besitzt. Nun lässt man das System relaxieren, bis sich ein Gleichgewichtszustand eingestellt hat.

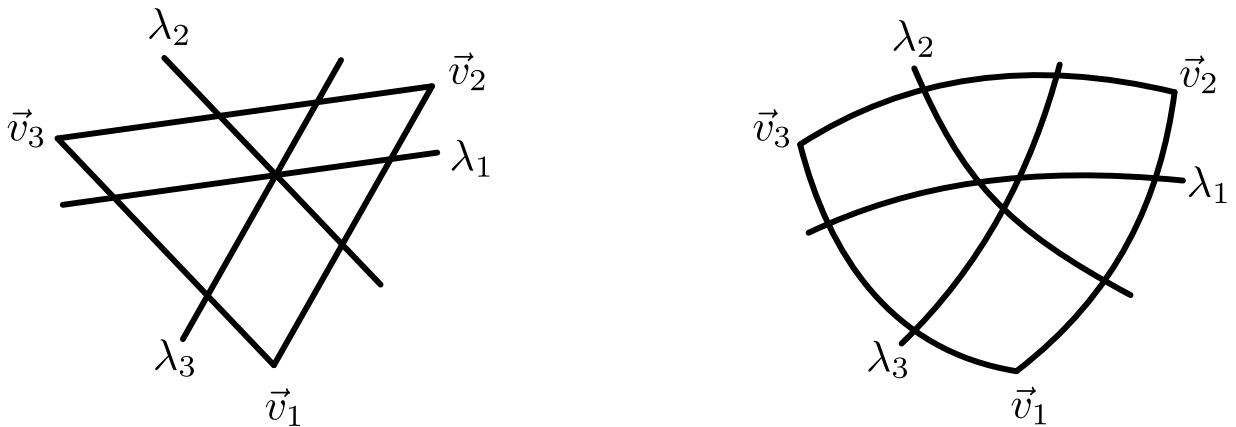
---

um sehr große Punkteverteilungen geht. Dieser Fall ist hier, da man an möglichst wenigen Elementen interessiert ist, nur bedingt anwendbar.

<sup>4</sup>Da hier eine variable Elementgrößenverteilung vorliegt, wird der maßgebliche Wert auf der Kantenmitte bezogen.

## A.2. Erzeugung der runden Elemente für die Kugel

Um die runden Elemente in höheren Ordnungen auch im 3D anwenden zu können, muss eine Darstellung der Randknoten existieren. Der in dieser Arbeit für die 3D-Projekte verwendete Netzgenerator NetGen (siehe [60]) liefert nur Oberflächendarstellung bis zur Ordnung 2, das heißt, dass die Oberfläche aus Paraboloidenabschnitten zusammengesetzt wird. Um nun eine höhere Auflösung zu erreichen wurde die Oberfläche, genauer die Projektion der Knoten von der Dreiecksfläche auf die Kugeloberfläche, mit nachfolgendem Verfahren bestimmt. Dabei ist zu beachten, dass es im Gegensatz zum 2D, wo eine reguläre Kurve der Bogenlänge nach immer auf eine Linie abgewickelt werden kann, keine isometrische Abbildung einer Kugeloberfläche auf eine Ebene existiert. Ausgegangen wird von



**Abbildung A.2.1.:** Übergang von der Punkteverteilung im Dreieck auf die Punkteverteilung auf der Kugeloberfläche

den drei Eckvertices ( $\vec{v}_i$ ) im Ursprungs-dreieck (siehe Abbildung A.2.1). Als nächstes werden die Dreieckskanten bestimmt. Dazu wurde hier die kürzeste Verbindung zwischen zwei Punkten auf der Kugel, der Großkreis, gewählt. Mit Hilfe diesem können nun die Kantenpunkte der Bogenlänge nach übertragen werden. Damit steht zumindest die Umrandung der projizierten Fläche fest. Für die Volumen-knoten bedient man sich der in Kapitel 3 vorgestellten baryzentrischen Koordinaten. So teilt  $\lambda_1$  die Strecken  $\vec{v}_1 - \vec{v}_2$  und  $\vec{v}_1 - \vec{v}_3$  der Kantenlänge nach. Überträgt man dieses Teilverhältniss auf die Kantenbögen auf der Kugel, hat man wieder zwei Zwischenpunkte, die sich mittels eines Großkreises verbinden lassen. Ebenso verfährt man mit den beiden anderen Koordinaten. Nun treffen sich auf der Kugeloberfläche die so gewonnenen Bögen im Allgemeinen nicht im selben Punkt. Daher wird für die Zielkoordinate der Mittelwert der drei Schnittpunkte gewählt. Abschließend muss noch, da der Mittelwert der drei Punkte nicht mehr auf der Kugeloberfläche liegt, der Abstand des neu konstruierten Punktes zum Kugelmittelpunkt korrigiert werden.

# Literaturverzeichnis

- [1] J. B. Andersen and V. V. Solodukhov. Field behavior near a dielectric wedge. *IEEE Transactions on Antennas and Propagation*, AP-26:598, 1978.
- [2] J.-P. Berenger. A perfectly matched layer for the absorption of electrodynamic waves. *Journal of Computational Physics*, 114:185, 1994.
- [3] N. Bloembergen and P. S. Pershan. Light waves at the boundary of nonlinear media. *Physical Review*, 128:606, 1962.
- [4] C. F. Bohren and D. R. Huffman. *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, New York, 1998.
- [5] F. Brown, R. E. Parks, and A. M. Sleeper. Nonlinear optical reflections from a metal boundary. *Physical Review Letters*, 14:1029, 1965.
- [6] G. W. Bryant, F. J. G. de Abajo, and J. Aizpurua. Mapping the plasmon resonances of metallic nanoantennas. *Nano Letters*, 8:631, 2008.
- [7] K. Busch, M. König, and J. Niegemann. Discontinuous galerkin methods in nanophotonics. *Laser and Photonics Reviews*, 5:773, 2011.
- [8] H. Cheng and P. B. Miller. Nonlinear optical theory in solids. *Physical Review A*, 134:683, 1964.
- [9] M. Corvi and W. L. Schaich. Hydrodynamic-model calculation of second-harmonic generation at a metal surface. *Physical Review B*, 33:3688, 1986.
- [10] F. J. G. de Abajo and A. Howie. Retarded field calculation of electron energy loss in inhomogeneous dielectrics. *Physical Review B*, 65:115418, 2002.
- [11] R. Diehl, K. Busch, and J. Niegemann. Comparison of low-storage runge-kutta schemes for discontinuous galerkin time-domain simulations of maxwell's equations. *Journal of Computational and Theoretical Nanoscience*, 7:1572, 2010.

- [12] J. Geisel, K.-H. Muth, and W. Heinrich. The behavior of electromagnetic fields at edges of media with finite conductivity. *IEEE Transactions on Microwave Theory and Techniques*, 40:158, 1992.
- [13] G. H. Golub and J. H. Welsch. Calculation of gauss quadrature rules. Technical Report CS 81, Computer Science Departement, School of Humanities and Sciences, Stanford University, 1967.
- [14] C. Hafner. *Post-modern Electromagnetics: Using Intelligent MaXwell Solvers*. John Wiley & Sons, New York, 1999.
- [15] J. S. Hesthaven and T. Warburton. Nodal high-order methods on unstructured grids: I. time-domain solution of maxwell’s equations. *Journal of Computational Physics*, 181:186, 2002.
- [16] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer, New York, 2008.
- [17] A. Hille, R. Kullock, S. Grafström, and L. M. Eng. Improving nano-optical simulations through curved elements implemented within the discontinuous galerkin method. *Journal of Computational and Theoretical Nanoscience*, 7:1581, 2010.
- [18] G. B. Jacobs and J. S. Hesthaven. High-order nodal discontinuous galerkin particle-in-cell method on unstructured grids. *Journal of Computational Physics*, 214:96, 2006.
- [19] S. S. Jha. Theory of optical harmonic generation at a metal surface. *Physical Review A*, 140:2020, 1965.
- [20] X. Ji, W. Cai, and P. Zhang. High-order dgtd methods for dispersive maxwell’s equations and modelling of silver nanowires. *International Journal for Numerical Methods in Engineering*, 69:308, 2007.
- [21] J. Jin. *The Finite Element Method in Electromagnetics*. John Wiley & Sons, New York, 1993.
- [22] P. B. Johnson and R. W. Christy. Optical constants of the noble metals. *Physical Review B*, 6:4370, 1972.
- [23] S. C. Kehr. *Scattering Scanning Near-Field Optical Microscopy an Anisotropic Dielectrics*. PhD thesis, Technische Universität Dresden, 2007.



- [24] A. M. Kern and O. J. F. Martin. Surface integral formulation for 3d simulations of plasmonic and high permittivity nanostructures. *Journal of the Optical Society of America A*, 26:732, 2009.
- [25] F. Kim, K. Sohn, J. Wu, and J. Huang. Chemical synthesis of gold nanowires in acidic solutions. *Journal of the American Chemical Society*, 130:14442, 2008.
- [26] A. Klöckner, T. Warburton, J. Bridge, and J. S. Hesthaven. Nodal discontinuous galerkin methods on graphic processors. *Journal of Computational Physics*, 228:7863, 2009.
- [27] M. König. *Discontinuous Galerkin methods in Nanooptics*. PhD thesis, Karlsruher Institut für Technologie, 2011.
- [28] M. König, K. Busch, and J. Niegemann. The discontinuous galerkin time-domain method for maxwell’s equations with anisotropic materials. *Photonics and Nanostructures - Fundamentals and Applications*, 8:303, 2010.
- [29] M. König, C. Prohm, K. Busch, and J. Niegemann. Stretched-coordinate pmls for maxwell’s equations in the discontinuous galerkin time-domain method. *Optics Express*, 19:4618, 2011.
- [30] E. Kretschmann and H. Raether. Radiative decay of non radiative surface plasmons excited by light (surface plasma waves excitation by light and decay into photons applied to nonradiative modes). *Zeitschrift für Naturforschung, Teil A*, 23:2135, 1968.
- [31] R. Kullock, A. Hille, A. Haußmann, S. Grafström, and L. M. Eng. Shg simulations of plasmonic nanoparticles using curved elements. *Optics Express*, 19:14426, 2011.
- [32] L. D. Landau and E. M. Lifschitz. *Lehrbuch der theoretischen Physik, Band IV Hydrodynamik*. Akademie Verlag, Berlin, 1991.
- [33] N. D. Lang and W. Kohn. Theory of metall surfaces: Charge density and surface energy. *Physical Review B*, 1:4555, 1970.
- [34] N. D. Lang and W. Kohn. Theory of metall surfaces: Work function. *Physical Review B*, 3:1215, 1971.
- [35] T. Lu, P. Zhang, and W. Cai. Discontinuous galerkin methods for dispersive and lossy maxwell’s equations and pml boundary conditions. *Journal of Computational Physics*, 200:549, 2004.

- [36] Y. Lui, C. D. Sarris, and G. V. Eleftheriades. Triangular-mesh-based fdtd analysis of two-dimensional plasmonic structures supporting backward waves at optical frequencies. *Journal of Lightwave Technology*, 25:938, 2007.
- [37] C. Matyssek, J. Niegemann, W. Hergert, and K. Busch. Computing electron energy loss spectra with the discontinuous galerkin time-domain method. *Photonics and Nanostructures - Fundamentals and Applications*, 9:367, 2011.
- [38] J. C. Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459, 1865.
- [39] J. A. Maytorena, W. L. Mochán, and B. S. Mendoza. Hydrodynamic model for second-harmonic generation at conductor surfaces with continuous profiles. *Physical Review B*, 51:2556, 1995.
- [40] J. A. Maytorena, W. L. Mochán, and B. S. Mendoza. Hydrodynamic model of second harmonic generation at metal surfaces. In *Proceedings of SPIE*, volume 2730, page 376, 1996.
- [41] J. Meixner. The behavior of electromagnetic fields at edges. *IEEE Transactions on Antennas and Propagation*, AP-20:442, 1972.
- [42] G. Mie. Beiträge zur optic trüber medien, speziell kolloidaler metallösungen. *Annalen der Physik*, 4:377, 1908.
- [43] D. L. Mills. *Nonlinear Optics*. Springer, New York, 1998.
- [44] F. Neubrech, T. Kolb, R. Lovrincic, G. Fahsold, A. Pucci, J. Aizpurua, T. W. Cornelius, M. E. Toimil-Molaes, R. Neumann, and S. Karim. Resonances of individual metal nanowires in the infrared. *Applied Physics Letters*, 89:253104, 2006.
- [45] J. Niegemann. *High-Order Methods for Solving Maxwell's Equations in the Time-Domain*. PhD thesis, Karlsruher Institut für Technologie, 2009.
- [46] J. Niegemann, M. König, K. Stannigel, and K. Busch. High-order time-domain methods for analysis of nano-phonic systems. *Photonics and Nanostructures - Fundamentals and Applications*, 7:2, 2009.
- [47] J. Niegemann, W. Pernice, and K. Busch. Simulation of optical resonators using dgtd and fdtd. *Journal of Optics A: Pure and Applied Optics*, 11:114015 (10), 2009.

- [48] Y. Nishikawa, T. Nagasawa, and K. Fujiwara. Silver island films for surface-enhanced infrared absorption spectroscopy: effect of island morphology on the absorption enhancement. *Vibrational Spectroscopy*, 6:43, 1993.
- [49] W. Nolting. *Grundkurs Theoretische Physik 3: Elektrodynamik*. Springer, New York, 7<sup>th</sup> edition, 2004.
- [50] L. Novotny. Effective wavelength scaling for optical antennas. *Physical Review Letters*, 98:266802, 2007.
- [51] M. Osawa. Surface-enhanced infrared absorption. *Topics in Applied Physics*, 81:163, 2001.
- [52] A. Otto. Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Zeitschrift für Physik*, 216:398, 1968.
- [53] P.-O. Persson and J. Peraire. Curved mesh generation and mesh refinement using lagrangian solid mechanics. In *Proceedings of the 47th AIAA Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, MIT, January 2009.
- [54] P.-O. Persson and G. Strang. A simple mesh generator in matlab. *SIAM Review*, 46:329, June 2004.
- [55] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computation*. Cambridge Universal Press, New York, 2007.
- [56] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical report, Los Alamos Scientific Laboratory, 1973.
- [57] P. Reichenbach. *Metallische Nanoantennen: Frequenzverdopplung und photochemische Reaktionen auf kleinen Skalen*. PhD thesis, Technische Universität Dresden, 2011.
- [58] J. Renger. *Excitation, Interaction, and Scattering of Localized and Propagating Surface Polaritons*. PhD thesis, Technische Universität Dresden, 2006.
- [59] J. Rudnick and E. A. Stern. Second-harmonic generation from metal surfaces. *Physical Review B*, 4:4274, 1971.
- [60] J. Schöberl. Netgen an advancing front 2d/3d-mesh generator based on abstract rules. *Computing and Visualization in Science*, 1:41, 1997.
- [61] J. E. Sipe, V. C. Y. So, and G. I. Stegeman. Analysis of second-harmonic generation at metal surfaces. *Physical Review B*, 21:4389, 1980.

- [62] J. Smajic, C. Hafner, L. Raguin, K. Tavzarashvili, and M. Mishrikey. Comparison of numerical methods for the analysis of plasmonic structures. *Journal of Computational and Theoretical Nanoscience*, 6:1, 2009.
- [63] K. Stannigel, M. König, J. Niegemann, and K. Busch. Discontinuous galerkin time-domain computations of metallic nanostructures. *Optics Express*, 17:14934, 2009.
- [64] A. Taflov. *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Boston, 2000.
- [65] A. Vial, A.-S. Grimault, D. Macías, D. Barchiesi, and M. L. de la Chapelle. Improved analytical fit of gold dispersion: Application to the modelin of extinction spectra with a finite-difference time-domain method. *Physical Review B*, 71:085416, 2005.
- [66] F. von Cube, S. Irsen, J. Niegemann, C. Matyssek, W. Hergert, K. Busch, and S. Linden. Spatio-spectral characterization of photonic meta-atoms with electron energy-loss spectroscopy. *Optical Material Express*, 1:1009, 2011.
- [67] T. Warburton. An explicit construction for interpolation nodes on the simplex. *Journal of Engineering Mathematics*, 56:247, April 2005.
- [68] M. Weber and A. Liebsch. Density-functional approach to second-harmonic generation at metal surfaces. *Physical Review B*, 35:7411, 1987.
- [69] M. T. Wenzel, T. Härtling, P. Olk, S. C. Kehr, S. Grafström, and L. M. Eng. Gold nanoparticle tips for optical field confinement in infrared scattering near-field optical microscopy. *Optics Express*, 16:12302, 2008.
- [70] S. M. Williams, A. D. Stafford, K. R. Rodriguez, T. M. Rogers, and J. V. Coe. Accessing surface plasmons with ni microarrays for enhanced ir absorption by monolayers. *Journal of Physical Chemistry B*, 107:11871, 2003.
- [71] H. C. Yee, N. D. Sandham, and M. J. Djomehri. Low-dissipative high-order shock-capturing methods using characteristic-based filters. *Journal of Computational Physics*, 150:199, 1999.
- [72] K. Yee. Numerical solution of initial boundary value problems involving maxwell’s equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:302, 1966.
- [73] Y. Zeng, W. Hoyer, J. Liu, S. W. Koch, and J. V. Moloney. Classical theory for second-harmonic generation from metallic nanoparticles. *Physical Review B*, 79:235109, 2009.

- [74] L. Zhan, T. Cui, and H. Liu. A set of symmetric quadrature rules on triangles and tetrahedra. *Journal of Computational Mathematics*, 27:89, 2009.



# Abbildungsverzeichnis

2.1.1	Brechung und Reflexion an einer ebenen Grenzfläche . . . . .	11
2.3.1	Dielektrische Funktion der verwendeten Materialien . . . . .	24
2.3.2	Dielektrische Funktion und LORENTZ-Fit von ZnPC . . . . .	25
2.3.3	Absorption am PML . . . . .	27
2.4.1	Dispersionsrelation des Oberflächenplasmons . . . . .	32
2.4.2	Streuung- und Absorption an einem Nanopartikel . . . . .	35
2.4.3	Wirkungsspektren und Nahfeldüberhöhung von einem Nanopartikel. . . .	36
2.4.4	Feldbilder von einem Nanopartikel . . . . .	37
2.4.5	Spektren und Nahfeld einer Gold-Nanoantenne . . . . .	37
2.6.1	Zeitevolution des $E$ -Feldes eines Silber-Partikels . . . . .	43
2.6.2	Vergleich zwischen Rand- und Volumendiskretisierung . . . . .	44
2.6.3	Schematischer Aufbau einer Simulation . . . . .	46
2.6.4	Pulsformen in Zeit- und Frequenzdomäne . . . . .	47
2.6.5	Zeitverlauf des DFT-Filters . . . . .	49
2.6.6	Zeitverlauf des DFT-Filters . . . . .	49
3.1.1	Räumliche Zerlegung eines Gebietes in Elemente . . . . .	54
3.1.2	Abbildung vom Referenzelement auf das physikalische Element . . . . .	60
3.1.3	Vergleich der Basisfunktionen im 1D . . . . .	61
3.1.4	Konstruktion der Knoten 1D . . . . .	63
3.1.5	Baryzentrische Koordinaten in 2D . . . . .	65
3.1.6	Konstruktion der Knoten in 2D . . . . .	67
3.1.7	Basisfunktionen der Ordnung 4 in 2D . . . . .	68
3.2.1	Bestimmung des CFL-Zeitschrittes im 1D . . . . .	79
3.2.2	Konvergenzanalyse für Lineare Elemente im 1D . . . . .	81
3.2.3	Skizze der Planparallele Platte . . . . .	82
3.2.4	Vergleich Materialmodelle im 1D . . . . .	83
3.2.5	Netz zum Berechnen des PML Fehlers im 2D . . . . .	84
3.2.6	Optimierung der PML Parameter im 2D . . . . .	85

3.3.1	Darstellung der Interpolationsfehler im 1D . . . . .	87
3.3.2	Skizze des Simulationsgebietes des isentropen Wirbels . . . . .	90
3.3.3	Darstellung der Dichte und Stromdichte des isentropen Wirbels . . . . .	91
3.3.4	Konvergenzanalyse des isentropen Wirbels im 2D . . . . .	92
4.0.1	Motivation der Runde Elemente . . . . .	93
4.1.1	Vorgehensweise zur Konstruktion der runden Elemente . . . . .	94
4.2.1	Netze 2D Kreiskavität . . . . .	98
4.2.2	Konvergenzanalyse der 2D-Kreiskavität . . . . .	99
4.3.1	Netze der 3D-Kugelkavität . . . . .	101
4.3.2	Beispiele verschiedener Oberflächenapproximationsordnungen . . . . .	101
4.3.3	Konvergenzanalyse der 3D-Kugelkavität . . . . .	102
5.1.1	Skizze der 2D-Kugel . . . . .	105
5.1.2	Serie uniform verfeinerter Netze der 2D-Kugel . . . . .	106
5.1.3	Serie selektiv verfeinerter Netze der 2D-Kugel . . . . .	106
5.1.4	Wirkungsspektren der 2D-Kugel für lineare Elemente 4.Ordnung . . . . .	107
5.1.5	Wirkungsspektren der 2D-Kugel für runde Elemente 4.Ordnung . . . . .	108
5.1.6	Feldüberhöhung der 2D-Kugel für lineare und runde Elemente . . . . .	109
5.1.7	Feldüberhöhung der 2D-Kugel bei 208 nm für runde Elemente . . . . .	111
5.1.8	Feldüberhöhung der 2D-Kugel bei 208 nm für lineare Elemente . . . . .	111
5.1.9	Spektren der 2D-Kugel bei selektive Verfeinerung 8.Ordnung . . . . .	112
5.1.10	Rechenzeiten für die 2D-Kugel . . . . .	113
5.1.11	Skizze eines V-Grabens . . . . .	115
5.1.12	Serie der V-Graben-Netze . . . . .	115
5.1.13	Spektren des V-Graben 4.Ordnung . . . . .	116
5.1.14	Feldüberhöhung des V-Graben $r = 2$ nm . . . . .	117
5.1.15	Feldüberhöhung des V-Graben $r = 10$ nm . . . . .	117
5.1.16	Rechenzeiten des V-Grabens . . . . .	118
5.2.1	Skizze der 3D-Kugel . . . . .	120
5.2.2	Ausschnitte der Netzes der 3D-Kugel . . . . .	121
5.2.3	Wirkungsquerschnitte der 3D-Kugel . . . . .	122
5.2.4	Feldüberhöhung der 3D-Kugel . . . . .	123
5.2.5	Nahfeld der 3D-Kugel mit $R = 1.5$ . . . . .	124
5.2.6	Nahfeld der 3D-Kugel mit $R > 1.5$ . . . . .	124
5.2.7	Skizze der flachen Antenne . . . . .	128
5.2.8	Netzausschnitt der flachen Antenne . . . . .	129



5.2.9	Spektren der flachen Antenne . . . . .	130
5.2.10	Feldstärke an Grenzfläche . . . . .	130
5.2.11	Fern- und Nahfeld Demodulation beim s-SNOM . . . . .	132
5.2.12	Skizze des 3D s-SNOM Systems . . . . .	133
5.2.13	Wirkungsquerschnitte des 3D s-SNOM . . . . .	135
5.2.14	Harmonische Analyse des abstandsabhängigen Streusignals . . . . .	137
5.2.15	Zeitabhängige Streuspektren des 3D s-SNOM . . . . .	139
5.3.1	Skizze der Simulation eines Nanopartikels . . . . .	142
5.3.2	Lineare und nichtlineare Wirkungsspektren einer Silberkugel . . . . .	143
5.3.3	Skizze des Dimer-Systems . . . . .	144
5.3.4	Wirkungsspektren des Dimers (Polarisation $\perp$ Dimerachse) . . . . .	145
5.3.5	Wirkungsspektren des Dimers (Polarisation $\parallel$ Dimerachse) . . . . .	147
A.1.1	Schritte der Netzerzeugung . . . . .	152
A.2.1	Projektion der Dreiecksfläche auf die Kugel . . . . .	154



# Tabellenverzeichnis

2.3.1	Verwendete Materialparameter von Silber und Gold . . . . .	24
2.3.2	LORENTZ-Fit von ZnPC . . . . .	25
3.1.1	Anzahl der Basisfunktionen in Abhängigkeit von Dimension und Ordnung	62
3.2.1	Randbedingungen für die MAXWELL-Gleichungen . . . . .	77
3.2.2	Lösung der Einheitswürfelkavität . . . . .	78
3.2.3	Verwendete Diskretisierung zur Bestimmung von $c_{CFL}$ . . . . .	78
3.2.4	CFL-Konstanten in Abhängigkeit von Dimension und Ordnung . . . . .	80
3.3.1	Randbedingungen für die EULER-Gleichung . . . . .	89
5.2.1	Normierte Rechenzeiten $t_{CPU}$ für die 3D-Kugel . . . . .	125



# Publikationsverzeichnis

## Artikel

- R. Kullock, W. R. Hendren, **A. Hille**, S. Grafström, P. R. Evans, R. J. Pollard, R. Atkinson and L. M. Eng „Polarization conversion through collective surface plasmons in metallic nanorod arrays“, *Opt. Express* **16**(26), 2008.
- T. Härtling, Y. Alaverdyan, **A. Hille**, M. T. Wenzel, M. Käll and L. M. Eng „Optically controlled interparticle distance tuning and welding of single gold nanoparticle pairs by photochemical metal deposition“, *Opt. Express* **16**(16), 2008.
- **A. Hille**, R. Kullock, S. Grafström and L. M. Eng „Improving Nano-Optical Simulations Through Curved Elements Implemented within the Discontinuous Galerkin Method“, *J. Comput. Theor. Nanosci.* **7**(8), 2010.
- R. Kullock, **A. Hille**, A. Haußmann, S. Grafström and L. M. Eng „SHG simulations of plasmonic nanoparticles using curved elements“, *Opt. Express* **19**(15), 2011
- T. Göhler, **A. Hille**, S. Grafström, R. Windisch, J. Moosburger and L. M. Eng „Enhancing the light emission on AlGaAs-based LEDs through single gold nanoparticles“, *J. Opt. Soc. Am. B* **28**(4), 2011.



# Danksagung

Da es nicht leicht ist, eine geeignete Reihenfolge bei der Danksagung zu finden, die niemanden verletzt oder bevorzugt, gilt dieser erste Satz als Dank all denjenigen, welche durch ihren Einsatz diese Arbeit möglich gemacht haben. Dennoch gebührt ein besonderer Dank einigen herausragenden Persönlichkeiten, die hier nicht unerwähnt bleiben sollen:

- Prof. Dr. Lukas M. Eng, für die Möglichkeit in seiner Arbeitsgruppe promovieren zu dürfen und die Freiheiten zu genießen, etwas neues ausprobieren zu können.
- Prof. Dr. Kurt Busch, für die enge Kooperation, die Möglichkeit ihn und seine Arbeitsgruppe kennenlernen zu dürfen und die freundliche Übernahme des Zweitgutachtens.
- Dr. Stefan Grafström, Dr. Alexander Haußmann und meiner Frau für die langjährige Unterstützung, die unzähligen Diskussionen und die Korrektur dieser Arbeit.
- Meiner Schwester, die von Physik vermutlich nie wieder etwas hören möchte.
- Meinem ehemaligen Betreuer René Kullock für die vielen Ratschläge und meinem Bürokollegen Tino Uhlig für das angenehme Arbeitsklima.
- Der ganzen SPM<sup>2</sup> Arbeitsgruppe am IAPP, für die gute Atmosphäre.
- Dr. Jens Niegemann und Dr. Michael König für die intensiven Gespräche rund um DG.
- Allen, die ich vergessen habe, hier zu erwähnen.

„Wo alles reibungslos läuft, findet man kein Glück.  
Glück ist da, wo die Katastrophe Pause macht.“  
– HORST EVERS AUS EVERSHORST





# Erklärung

Diese Dissertation wurde am Institut für Angewandte Physik/Photophysik der Fakultät Mathematik und Naturwissenschaften an der Technischen Universität Dresden unter wissenschaftlicher Betreuung von Prof. Dr. Lukas Eng angefertigt.

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere weiterhin, dass keinerlei frühere Promotionsverfahren stattgefunden haben.

Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften an der Technischen Universität Dresden in der Fassung vom 23.02.2011 an.

Andreas Hille  
Dresden, September 2012